

CÁLCULO DIFERENCIAL

en varias variables

Rubén Becerril Fonseca
Daniel R. Jardón Arcos
J. Guadalupe Reyes Victoria

UNIVERSIDAD AUTONOMA METROPOLITANA-IZTAPALAPA
Casa abierta al tiempo
División de Ciencias Biológicas y de la Salud



ISBN 970-31-0096-1

Primera Edición: 2002
Universidad Autónoma Metropolitana, Unidad Iztapalapa
Av. San Rafael Atlixco 186, Col. Vicentina
México, D.F. 09340

Impreso y hecho en México



CÁLCULO DIFERENCIAL
EN VARIAS VARIABLES

Rubén Becerril Fonseca
Daniel Jardón Arcos
J. Guadalupe Reyes Victoria

Departamento de Matemáticas
UAM-IZTAPALAPA


Casa abierta al tiempo



Prefacio

Uno de los principales problemas que tiene un lector estudioso de las ciencias básicas (física, química, matemáticas) o de la ingeniería, es el de encontrar aislados en la abundante literatura muchos temas de su interés. Por un lado, los textos clásicos orientados hacia las partes aplicadas de la ciencia que tienen relevancia, no son del todo accesibles al joven lector, y para su lectura imponen una cantidad considerable de prerrequisitos. Por otra parte, la literatura teórica en muchas ocasiones causa tedio en los aspirantes al ejercicio práctico de la ciencia.

La presente obra trata de equilibrar las dos cuestiones: la parte teórica de una forma simple, y su uso en las partes aplicadas de la ciencia, pensando en la formación del futuro científico, ingeniero, técnico, etcétera. Para su lectura se presuponen conocimientos elementales de cálculo diferencial de una variable. Los demás conceptos el lector no matemático puede irlos aprendiendo durante el camino.

En el primer capítulo hacemos un bosquejo de los elementos necesarios para leer el trabajo, como son los sistemas de ecuaciones, las matrices y determinantes. De esta manera se hace una reseña de los elementos básicos del **Álgebra Lineal**.

El capítulo 2 trata de los aspectos básicos de los objetos geométricos elementales en el espacio Euclidiano: las **rectas**, los **planos**. Para su construcción necesitamos los conceptos de **vector**, **ángulo** y **distancia**. Aquí se estudia también el problema de vectores y valores propios de una matriz cuadrada real.

En el capítulo 3 se estudian los elementos básicos de las **curvas** planas y espaciales, sus propiedades diferenciables: su velocidad y aceleración.

El capítulo 4 muestra el estudio de los campos escalares diferenciables $\mathbb{R}^n \to \mathbb{R}$, y los elementos para realizarlo: las **derivadas parciales**, el **gradiente**, el polinomio de Taylor, etcétera.



En el capítulo 5 se hace una muestra somera de la teoría básica de los campos vectoriales diferenciables del tipo $\mathbb{R}^n \to \mathbb{R}^m$ $(n, m \leq 3)$ y los conceptos asociados más importantes se enuncian y ejemplifican: **la divergencia, el rotor** y el **gradiente**.

En los apéndices se incluyen tópicos clásicos de los cursos de cálculo en varias variables como: elementos básicos de **superficies** en $\mathbb{R}^3$, orientación y **longitud** de **una curva** en $\mathbb{R}^3$, **límites** y **continuidad** de campos escalares, **planos tangentes, valores extremos** de un campo escalar y las **funciones implícitas**.

Este libro es producto de varios cursos de Matemáticas IV para los estudiantes de Ciencias Biológicas y de la Salud (CBS) en la Universidad Autónoma Metropolitana Iztapalapa, durante los años 1992 - 1999. La presentación es diferente de la de los cursos clásicos debido a que las necesidades de las propias licenciaturas (ingenieros bioquímicos, biotecnólogos y en alimentos) así lo requieren.

Deseamos manifestar nuestro agradecimiento al Dr. Gerardo Saucedo, Director de la División de CBS, al M. en C. Arturo Preciado, Secretario Académico de la División de CBS, a la Dra. María José Arroyo, ex-Directora de la División de CBI y al Dr. Ernesto Pérez, Jefe del Departamento de Matemáticas por todo el apoyo y entusiasmo que nos brindaron. También queremos resaltar la contribución de los profesores y alumnos que usaron versiones preliminares y cuyos valiosos comentarios nos ayudaron a mejorar el texto. La presentación final se logró gracias a la colaboración de Daniel Espinosa (Flash).

Por último, quisieramos agradecer a nuestras respectivas familias por toda la paciencia infatigable a lo largo de este proyecto.

R.B.F., D.R.J.A., J.G.R.V.
IZTAPALAPA 2002



Contenido


Casa abierta al tiempo



Capítulo 1

Sistemas de ecuaciones, matrices y determinantes

1.1 Sistemas de ecuaciones

Iniciamos con el problema de resolver un sistema de ecuaciones lineales de 2×2 con coeficientes reales de la forma

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases}$$

donde a_{11}, a_{12}, a_{21}, a_{22}, b_1 y b_2 son constantes reales y x, y son incógnitas.

$\triangleleft$ Despejamos a la variable x de la primera ecuación y la sustituimos en la segunda.

De la relación

$$a_{11}x + a_{12}y = b_1$$

se obtiene que

$$a_{11}x = b_1 - a_{12}y$$

lo que implica que, si $a_{11} \neq 0$, entonces

$$x = \frac{b_1 - a_{12}y}{a_{11}}$$

Al sustituir en la segunda ecuación del sistema se tiene

$$b_2 = a_{21}\left(\frac{b_1 - a_{12}y}{a_{11}}\right) + a_{22}y \qquad (*)$$

que nos permite resolver para la variable y de la siguiente forma.

De la ecuación

$$b_2 = \frac{a_{21}(b_1 - a_{12}y) + a_{11}a_{22}y}{a_{11}}$$

se obtiene

$$a_{11}b_2 = a_{21}b_1 - a_{21}a_{12}y + a_{11}a_{22}y$$

de donde,

$$a_{11}b_2 - a_{21}b_1 = y(a_{11}a_{22} - a_{21}a_{12})$$

De esta manera, si

$$a_{11}a_{22} - a_{21}a_{12} \neq 0$$

entonces podemos despejar a y, quedando

$$\frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}} = y$$

Es decir,

$$y = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}$$

Al sustituir esta igualdad en la ecuación $(*)$ se obtiene la indeterminada x de la siguiente cadena de igualdades

$$x = \frac{b_1 - a_{12}\left(\frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}\right)}{a_{11}} = \frac{b_1 - \frac{a_{12}(a_{11}b_{12} - a_{21}b_1)}{a_{11}a_{22} - a_{21} - a_{12}}}{a_{11}}$$

$$= \frac{\frac{b_1(a_{11}a_{22} - a_{21}a_{12}) - a_{12}(a_{11}b_2 - a_{21}b_1)}{a_{11}a_{22} - a_{21}a_{12}}}{a_{11}} = \frac{a_{11}(b_1a_{22} - a_{12}b_2)}{a_{11}(a_{11}a_{22} - a_{21}a_{21})}$$

De donde,

$$x = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{21}a_{12}} \quad \triangleright$$

Concluimos la discusión con el siguiente lema.

LEMA 1.1 *Para el sistema lineal de dos ecuaciones con dos incógnitas* x, y

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_y \end{cases}$$

con a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2 *números reales, se tiene la solución dada por la pareja*

$$x = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{21}a_{12}}$$



$$y = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}$$

sabiendo que $a_{11}a_{22} - a_{21}a_{12} \neq 0$

NOTACIÓN. Para una expresión real $ad - bc$ definimos el arreglo

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

mediante la igualdad

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

EJEMPLO. Para a_{11}, a_{12}, a_{12}, , a_{22}, b_1, b_2 números reales arbitrarios tenemos

$$\triangleleft \quad a_{11}a_{21} - a_{21}a_{12} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$b_1a_{22} - a_{12}b_2 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}$$

$$a_{11}b_2 - a_{21}b_1 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} \quad \triangleright$$

En esta notación, las soluciones del sistema dado en el Lema 1.1 se escriben

$$x = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}$$

Para tal sistema de ecuaciones se introduce el siguiente arreglo formado por los coeficientes que intervienen en el sistema

$$\begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{pmatrix}$$

y le llamamos la **matriz principal** del sistema.

Se obtienen de esta las **submatrices** cuadradas (2×2),

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}, \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}$$

y a cada arreglo se le asocia un número distinguido propio

$$D = \left| \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$



$$D_x = \left|\begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}\right| = \begin{vmatrix} b_1 & a_{12} \\ a_2 & a_{22} \end{vmatrix} = b_1 a_{22} - a_{12} b_2$$

$$D_y = \left|\begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}\right| = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} = a_{11} b_2 - b_1 a_{21}$$

llamados sus **determinantes** correspondientes.

Con esta última notación se tiene el siguiente resultado.

COROLARIO 1.1 *Las soluciones x, y del sistema inicial se calculan por*

$$x = \frac{D_x}{D} \quad y = \frac{D_y}{D}$$

sabiendo que $D \neq 0$.

EJEMPLO. Resolver el sistema

$$\begin{cases} 2x - 3y = 1 \\ 3x + 2y = 2 \end{cases}$$

$\lhd$ La matriz principal del sistema es

$$\begin{pmatrix} 2 & -3 & 1 \\ 3 & 2 & 2 \end{pmatrix}$$

Las submatrices asociadas son

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & -3 \\ 2 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$$

y los determinantes correspondientes se calculan

$$D = \left|\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}\right| == \begin{vmatrix} 2 & -3 \\ 3 & 2 \end{vmatrix} = 4 + 9 = 13 \neq 0$$

$$D_x = \left|\begin{pmatrix} 1 & -3 \\ 2 & 2 \end{pmatrix}\right| = \begin{vmatrix} 1 & -3 \\ 2 & 2 \end{vmatrix} = 2 + .6 = 8$$

$$D_y = \left|\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}\right| = \begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} = 4 - 3 = 1$$

Esto nos lleva a que las soluciones del sistema de ecuaciones son

$$x = \frac{D_x}{D} = \frac{8}{13}$$

$$y = \frac{D_y}{D} = \frac{1}{13} \quad \rhd$$

En la discusión mostrada al resolver un sistema lineal de dos ecuaciones con dos incógnitas nos encontramos con los conceptos auxiliares de solución que son, las **matrices** y los **determinantes**. Estudiamos estos objetos matemáticos y sus propiedades en las próximas secciones. Más adelante volveremos al problema de resolver sistemas de ecuaciones lineales.

1.2 Matrices

Un arreglo de la forma

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \cdots a_{1n} \\ a_{21} & a_{22} & a_{23} \cdots a_{2n} \\ \vdots & & \\ a_{m1} & a_{m2} & a_{m3} \cdots a_{mn} \end{pmatrix}$$

donde cada a_{ij} es un número real, se llamará una **matriz real** de **dimensiones** $m \times n$. Esto es, el arreglo consta de $m -$ renglones y $n -$ columnas

El i-ésimo renglón de la matriz sería

$$(a_{i1}, a_{i2}, a_{i3} \cdots a_{in})$$

y se llamaría el i-ésimo **vector renglón.**

Al arreglo

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

se le denomina el j-ésimo **vector columna.**

Una **entrada** a_{ij} de la matriz estaría en el i-ésimo renglón y la j-ésima columna (i define los renglones y j define las columnas).

EJEMPLO. La matriz real

$$\begin{pmatrix} 1 & 3 & 4 & 6 & 8 \\ 2 & 3 & 2 & -1 & 0 \\ 4 & 3 & 5 & 7 & -2 \end{pmatrix}$$

es de (3×5).

◁ Mencionamos algunos elementos de la matriz

$$a_{34} = 7, \quad a_{15} = 8, \quad a_{23} = 1, \quad a_{32} = 3$$

El segundo vector renglón es

$$(2 \quad 3 \quad 1 \quad -1 \quad 0)$$

mientras que el tercer vector columna es

$$\begin{pmatrix} 4 \\ 2 \\ 5 \end{pmatrix} \quad \triangleright$$



Una matriz con dimensiones $m = n$ tal se llamará **matriz cuadrada**.

EJEMPLO. Las matrices

$$\begin{pmatrix} 1 & 2 \\ 4 & -2 \end{pmatrix}_{2\times 2}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 3 \\ 0 & 1 & 0 \end{pmatrix}_{3\times 3}$$

son cuadradas.

A una matriz con entradas idénticamente nulas se llamará **matriz nula** y se denota por

$$O = \begin{pmatrix} 0 & 0 & \cdots 0 \\ 0 & 0 & \cdots 0 \\ \vdots & & \\ 0 & 0 & \cdots 0 \end{pmatrix}_{m\times n}$$

1.3 Operaciones básicas de matrices

En adelante identificaremos a las matrices reales por letras mayúsculas:

$$A, B, C, \cdots M, \cdots$$

y sus entradas reales por minúsculas

$$a_{ij}, b_{ij}, c_{ij}, \cdots m_{ij}, \cdots$$

Dos matrices se pueden sumar si tienen las mismas dimensiones, es decir, si

$$A = A_{m\times n}, \quad B = B_{m\times n}$$

entonces $A + B$ tiene sentido. De hecho, si las matrices tienen entradas

$$A = (a_{ij}), \quad B = (b_{ij})$$

entonces su suma se define por la suma de entradas correspondientes, esto es,

$$A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

EJEMPLO. Sean las matrices reales

$$A = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 1 \\ -1 & 4 & 2 \end{pmatrix}_{3\times 3}$$



$$B = \begin{pmatrix} 2 & 3 & 4 \\ -1 & 4 & 2 \end{pmatrix}_{2\times 3}$$

$\lhd$ No tienen sentido la suma matricial $A + B$ por ser de diferentes dimensiones. $\rhd$

EJEMPLO. Sean las matrices reales

$$A = \begin{pmatrix} -1 & 4 & 3 & 2 \\ -2 & 1 & 8 & 2 \end{pmatrix}_{2\times 4}$$

$$B = \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 4 & -1 & 2 \end{pmatrix}_{2\times 4}$$

$\lhd$ Su suma tiene sentido y se calcula

$$A+B = \begin{pmatrix} -1 & 4 & 3 & 2 \\ -2 & 1 & 8 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 4 & -1 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 5 & 3 & 3 \\ -3 & 5 & 7 & 4 \end{pmatrix} \rhd$$

Una cantidad **escalar** es un número real arbitrario y lo identificamos por las letras griegas

$$\lambda,\ \mu,\ \tau, \cdots$$

Dada la matriz $A = (a_{ij})_{m\times n}$, se define la nueva matriz λA de $m \times n$ obtenida al multiplicar por el escalar λ, como aquella que se obtiene de multiplicar por λ cada una de sus entradas, esto es,

$$\lambda A = \lambda(a_{ij}) = (\lambda a_{ij})$$

EJEMPLOS.

i. Si tomamos

$$\lambda = 3, \quad A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ -1 & 4 & 2 \end{pmatrix}$$

entonces

$$\lambda A = 3 \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 1 \\ -1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 0 & -9 \\ 0 & 3 & 3 \\ -3 & 12 & 6 \end{pmatrix}$$

ii. Al considerar

$$\mu = -1, \quad B = \begin{pmatrix} 2 & 3 & 0 \\ -1 & 4 & 2 \end{pmatrix}$$

se obtiene la matriz

$$\mu B = -1 \begin{pmatrix} 2 & 3 & 0 \\ -1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} -2 & -3 & 0 \\ 1 & -4 & -2 \end{pmatrix}$$



OBSERVACIONES.

1. La matriz $O = (0)$ (**nula**) es el elemento matricial que al sumarse a cualquier matriz no la altera:

$$A + O = A$$

2. Dada A una matriz arbitraria y $\lambda = -1$, se define $-A = (-1)A$ como la matriz inversa aditiva de A, tal que

$$A + (-1)A = O$$

EJEMPLO. Sean las matrices

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 1 \\ 0 & -3 \end{pmatrix}$$

Calcular: $A + B$, $3A$, $A + 2B$, $A - B$.

$\triangleleft$ Mediante calculos directos, obtenemos,

$$A + B = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} -1 & 1 \\ 0 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 2 & -1 \end{pmatrix}$$

$$3A = 3\begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -3 \\ 6 & 6 \end{pmatrix}$$

$$A + 2B = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix} + 2\begin{pmatrix} -1 & 1 \\ 0 & -3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} -2 & 2 \\ 0 & -6 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 2 & -4 \end{pmatrix}$$

$$A - B = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix} - \begin{pmatrix} -1 & 1 \\ 0 & -3 \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ 2 & 5 \end{pmatrix} \quad \triangleright$$

Sea $A = (a_{ij})$ una matriz arbitraria de $m \times n$, se define la **matriz transpuesta** de A como la matriz de $n \times m$ que se obtiene de a_{ij} intercambiando renglones por columnas. Identifitada por A^t, sus entradas la definen por

$$A^t = (a_{ji})$$

EJEMPLOS.

i. Dada la matriz (2×3)

$$A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 2 \end{pmatrix}_{2\times 3}$$



se tiene que su transpuesta es la matriz (3×2)

$$A^t = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}_{3\times 2}$$

ii. Para la matriz renglón

$$A = (1 \quad -1 \quad 4)$$

se obtiene la matriz transpuesta dada por el vector columna,

$$A^t = \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix}$$

iii. Para la matriz cuadrada de 2×2

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$$

su transpuesta es

$$A^t = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$$

Sean $A_{m\times n}, B_{n\times s}$ dos matrices reales. Se define su producto como la matriz $C_{m\times s}$ obtenida de la siguiente forma:

$$A_{m\times n}B_{m\times s} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1s} \\ b_{21} & b_{22} & \cdots & b_{2s} \\ & \vdots & & \\ b_{n1} & b_{n2} & \cdots & b_{ns} \end{pmatrix}$$

$$= \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1s} \\ c_{21} & c_{22} & \cdots & c_{2s} \\ & \vdots & & \\ c_{m1} & c_{m2} & \cdots & c_{ms} \end{pmatrix} = (c_{ij}) = C_{m\times s}$$

donde la entrada c_{ij} de C se obtiene de multiplicar los elementos del i-éisimo renglón de A con los de la j-ésima columna de B, y realizar la suma de estos productos.

Esto es,

$$c_{ij} = \underbrace{\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}}_{i\text{-ésimo renglón}} \underbrace{\begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}}_{j\text{-ésima columna}}$$

$$= a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

siendo que el producto del i-ésimo renglón de A con la j-ésima columna de B se define por la suma anterior.

Si tenemos

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix}$$

donde

$$A_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}$$

y

$$B = \begin{pmatrix} B^1 & B^2 & \cdots & B^s \end{pmatrix}$$

siendo el vector columna

$$B^j = \begin{pmatrix} b_{ij} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix}$$

entonces los elementos del producto se calculan, mediante

$$c_{ij} = A_i B^j = \sum_{k=1}^{n} a_{ik}b_{kj}$$

con el producto estipulado por la suma anterior.

En otras palabras, para poder multiplicar dos matrices $A = A_{m\times n}$ y $B = B_{n\times s}$ es necesario que el número n de columnas de A coincida con el número n de renglones de B, y el resultado (producto) es una matriz $C = C_{m\times s}$ de dimensiones $m \times s$



EJEMPLO. Sean las matrices

$$A = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}_{3\times 2}$$

$$B = \begin{pmatrix} 1 & 0 & 1 & -1 & 4 \\ -1 & 2 & 3 & -1 & 0 \end{pmatrix}_{2\times 3}$$

Calcular BA y AB.

$\lhd$ **a.** $BA = B_{2\times 5}\ A_{3\times 2}$, como $5 \neq 3$ no tiene sentido el producto.

b. $AB = A_{3\times 2}\ B_{2\times 5}$, en este caso tiene sentido el producto y se realiza

$$\begin{pmatrix} 3 & 2 \\ -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 & 4 \\ -1 & 2 & 3 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 9 & -5 & 12 \\ -1 & 0 & -1 & 1 & -4 \\ -1 & 2 & 3 & -1 & 0 \end{pmatrix}$$

A continuación mostramos el cálculo de algunos elementos c_{ij} del producto de las matrices dadas

$$c_{11} = (3 \quad 2)(1 \quad -1) = 3 - 2 = 1$$

$$c_{12} = (3 \quad 2)(0 \quad 2) = 0 + 4 = 4$$

$$c_{13} = (3 \quad 2)(1 \quad 3) = 3 + 6 = 9$$

$$c_{14} = (3 \quad 2)(-1 \quad -1) = -3 - 2 = -5$$

$$c_{15} = (3 \quad 2)(4 \quad 0) = 12 + 0 = 12$$

$$c_{23} = (-1 \quad 0)(0 \quad 2) = 0 + 2 = 2$$

$$c_{35} = (0 \quad 1)(4 \quad 0) = 0$$

La matriz obtenida es de (3×5) $\rhd$

EJEMPLO. Sean las matrices

$$A = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 3 & 2 \end{pmatrix}_{2\times 3} \qquad B = \begin{pmatrix} 3 & 4 \\ -1 & 2 \\ 2 & 1 \end{pmatrix}_{3\times 2}$$

Calcular: AB y BA.

$\lhd$ **a.** $AB = A_{2\times 3}B_{3\times 2} = C_{2\times 2}$ es una matriz cuadrada de (2×2)

$$AB = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ -1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 15 & 15 \\ 4 & 12 \end{pmatrix}$$



los cálculos se muestran a continuación:

$$\begin{bmatrix} c_{11} = 6 - 1 + 10 = 15 & c_{21} = 3 - 3 + 4 = 4 \\ c_{12} = 8 + 2 + 5 = 15 & c_{22} = 4 + 6 + 2 = 12 \end{bmatrix}$$

b. $BA = B_{3\times 2}A_{2\times 3} = C_{3\times 3}$ es una matriz cuadrada de (3×3).

$$BA = \begin{pmatrix} 3 & 4 \\ -1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 5 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 15 & 23 \\ 0 & 5 & -1 \\ 5 & 5 & 12 \end{pmatrix}$$

Dejamos al lector la verificación de los cálculos. $\rhd$

El ejemplo precedente muestra que en general el producto de matrices no es conmutativo, es decir, en general $AB \neq BA$

TEOREMA 1.1 *Sean las matrices* $A = A_{m\times n}$, $B = B_{m\times n}$, $C = C_{\ell\times m}$

a. *Se tiene la siguiente regla de distribución del producto con la suma de matrices*

$$C(A + B) = CA + CB$$

b. *Si* λ *es un número real (***escalar***) entonces*

$$C(\lambda B) = (\lambda C)B$$

$\lhd$ Damos a la matriz $(A + B) = (A^j + B^j)$ en forma de columnas.

Por otro lado, si escribimos a la matriz $C = (C_i)$ en renglones entonces

$$C(A + B) = (C_i)\,(A^j + B^j) = (C_iA^j) + (C_i\,B^j) = CA + CB \quad \rhd$$

TEOREMA 1.2 *Si* $A = A_{m\times n}$, $B = B_{n\times s}$, $C = C_{\ell\times m}$ *son matrices reales, entonces el producto es asociativo, esto es,*

$$C(A\,B) = (CA)B$$

$\lhd$ Es un cálculo directo y se omite $\rhd$

EJEMPLO. Sean las matrices.

$$A = \begin{pmatrix} .2 & 1 \\ -1 & 0 \end{pmatrix}_{(2\times 2)} \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}_{(2\times 2)} \quad C = \begin{pmatrix} 3 & 2 \\ 2 & 1 \\ 1 & 0 \end{pmatrix}_{3\times 2}$$

a. Verificamos que se cumple la relación

$$C(A + B) = (CA + CB)$$



◁ Primeramente realizamos la suma de matrices $A+B$ y luego calculamos $C(A+B)$

$$C(A+B)=\begin{pmatrix}3&2\\2&1\\1&0\end{pmatrix}\left[\begin{pmatrix}2&1\\-1&0\end{pmatrix}+\begin{pmatrix}0&1\\-1&0\end{pmatrix}\right]$$

$$=\begin{pmatrix}3&2\\2&1\\1&0\end{pmatrix}\begin{pmatrix}2&2\\-2&0\end{pmatrix}=\begin{pmatrix}2&6\\2&1\\2&2\end{pmatrix}$$

Por otro lado,

$$CA+CB=\begin{pmatrix}3&2\\2&1\\1&0\end{pmatrix}\begin{pmatrix}2&1\\-1&0\end{pmatrix}+\begin{pmatrix}3&2\\2&1\\1&0\end{pmatrix}\begin{pmatrix}0&1\\-1&0\end{pmatrix}$$

$$=\begin{pmatrix}4&3\\3&2\\2&1\end{pmatrix}+\begin{pmatrix}-2&3\\-1&2\\0&1\end{pmatrix}=\begin{pmatrix}2&6\\2&1\\2&2\end{pmatrix}$$

lo que prueba la validez de la relación. ▷

b. Ahora comprobamos la asociatividad

$$C(AB)=(CA)B$$

◁ Por un lado,

$$C(AB)=\begin{pmatrix}3&2\\2&1\\1&0\end{pmatrix}\left[\begin{pmatrix}2&1\\-1&0\end{pmatrix}\begin{pmatrix}0&1\\-1&0\end{pmatrix}\right]$$

$$=\begin{pmatrix}3&2\\2&1\\1&0\end{pmatrix}\begin{pmatrix}-1&2\\0&-1\end{pmatrix}=\begin{pmatrix}-3&1\\-2&3\\-1&2\end{pmatrix}$$

Por otro lado,

$$(CA)B=\left[\begin{pmatrix}3&2\\2&1\\1&0\end{pmatrix}\begin{pmatrix}2&1\\-1&0\end{pmatrix}\right]\begin{pmatrix}0&1\\-1&0\end{pmatrix}$$

$$=\begin{pmatrix}4&3\\3&2\\2&1\end{pmatrix}\begin{pmatrix}0&1\\-1&0\end{pmatrix}=\begin{pmatrix}-3&1\\-2&3\\-1&2\end{pmatrix}$$



lo que prueba la igualdad. ▷

La matriz cuadrada

$$I = \begin{pmatrix} 1 & 0 & 0 \cdots 0 \\ 0 & 1 & 0 \cdots 0 \\ 0 & 0 & 1 \cdots 0 \\ \vdots & & \\ 0 & 0 & 0 \cdots 1 \end{pmatrix}_{n \times n}$$

con $1's$ en la diagonal y los otros elementos nulos, es tal que si $A = A_{m \times n}$ es una matriz arbitraria, entonces

$$AI = A$$

I se llamará **matriz identidad** de $n \times m$.

EJEMPLO. Sea la matriz

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 1 \\ 1 & 0 \end{pmatrix}_{3 \times 2}$$

y sea $I = I_{2 \times 2}$ la matriz identidad, se tiene que

$$\triangleleft \quad AI = AI_{2 \times 2} = \begin{pmatrix} 3 & 2 \\ 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 2 & 1 \\ 1 & 0 \end{pmatrix} = A \quad \triangleright$$

1.4 Determinantes de orden tres

En esta parte definimos para las matrices cuadradas de dimensiones 2×2 y 3×3, un número característico: su **determinante**. Este concepto, así como sus propiedades, pueden ser generalizadas para matrices cuadradas de dimensiones mayores.

Para una matriz cuadrada de 2×2 real

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

se define el número característico $|A|$ llamado su **determinante**, por la igualdad

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

Cambiando un poco la notación, lo anterior nos dice que para la matriz

$$A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

su determinante se calcula por la igualdad

$$|A| = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1 = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = |A^t|$$

Esto nos dice que los determinantes de una matriz A de orden 2×2 y la de su transpuesta A^t son iguales.

EJEMPLO. El determinante de la matriz

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

se calcula

$$\lhd \quad |A| = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 2 - 1 = 1 \quad \rhd$$

Para una matriz real de 3×3

$$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

definimos de igual forma su determinante $|A|$ como

$$|A| = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

donde los determinantes 2×2 en el desarrollo se calculan ordinariamente.

EJEMPLO. Calculamos el determinante de la matriz

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$$

$$\lhd \quad \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix} = 1 \begin{vmatrix} 3 & 4 \\ 4 & 5 \end{vmatrix} - 2 \begin{vmatrix} 2 & 4 \\ 3 & 5 \end{vmatrix} + 3 \begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix}$$

$$= 1(15 - 16) - 2(10 - 12) + 3(8 - 9) = -1 + 4 - 3 = 0 \quad \rhd$$

Dado el determinante

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

los números a_i, b_i, c_i, $(i = 1, 2, 3)$ se llaman los **elementos del determinante** .

Si se considera un elemento del determinante, se dice que el determinante de la matriz 2×2 que se obtiene al quitar el renglón y la columna donde se localiza tal elemento es su **menor** correspondiente.

EJEMPLO. En el determinante

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$$

para el número 3 en el centro, su menor correspondiente es

$$\begin{vmatrix} 1 & 3 \\ 3 & 5 \end{vmatrix}$$

EJEMPLO. En el determinante

$$\begin{vmatrix} 0 & 1 & 6 \\ 4 & -11 & 14 \\ \pi/2 & \sqrt{2} & 1 \end{vmatrix}$$

para $\pi/2$, su menor es

$$\begin{vmatrix} 1 & 6 \\ -11 & 14 \end{vmatrix}$$

Definimos la **paridad** de un elemento de un determinante como la suma del número, el renglón y del número de la columna donde se encuentra.

EJEMPLO. En el determinante

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & \circ 4 & 5 \end{vmatrix}$$

el elemento indicado 4 tiene paridad$= 3 + 2 = 5$.

Es el mismo determinante,

$$\begin{vmatrix} 1 & \circ 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$$



el elemento 2 tiene paridad$= 1 + 2 = 3$.

Igualmente. en

$$\begin{vmatrix} 1 & 2 & \circ 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$$

el elemento 3 tiene la paridad$= 1 + 3 = 4$

Definimos para un elemento a_{ij} de un determinante de 3×3 su **cofactor** A_{ij} como el número (que lleva signos)

$$A_{ij} = (-1)^{i+j} M_{ij}$$

donde i es el número de renglón, j el número de columna y M_{ij} es su menor correspondiente.

EJEMPLO. Consideremos nuevamente el determinante del ejemplo anterior, entonces,

i. El elemento 4 indicado,

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & \circ 4 & 5 \end{vmatrix}$$

tiene cofactor $(-1)^{3+2} \begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} = - \begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} = -(4-6) = 2$

ii. Análogamente, el elemento 1 indicado,

$$\begin{vmatrix} \circ 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$$

tiene cofactor $(-1)^{1+1} \begin{vmatrix} 3 & 4 \\ 4 & 5 \end{vmatrix} = \begin{vmatrix} 3 & 4 \\ 4 & 5 \end{vmatrix} = 15 - 16 = -1$

iii. De igual manera, en el mismo determinante,

$$\begin{vmatrix} 1 & \circ 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$$

2 tiene cofactor $A = (-1)^{2+1} \begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = - \begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = -(10-12) = 2$

Observemos que inicialmente un determinante se desarrolló

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$



Si se definen

$$A_1 = \text{cofactor } (a_1) = (-1)^{1+1}\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} = \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}$$

$$B_1 = \text{cofactor } (b_1) = (-1)^{1+2}\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} = -\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}$$

$$C_1 = \text{cofactor } (c_1) = (-1)^{1+3}\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} = \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

en esta notación, el determinante se puede desarrollar mediante

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1A_1 + b_1B_1 + c_1C_1$$

Un resulado más general sobre el desarrollo de un determinante viene dado por el siguiente teorema, cuya prueba omitimos.

TEOREMA 1.3 (del desarrollo) *Cualquier determinante* 3×3 *se calcula como la suma de los términos de un renglón (o una columna) por sus cofactores.*

◁ En particular, al desarrollar por la tercer columna,

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = c_1C_1 + c_2C_2 + c_3C_3$$

donde el cofactor de c_1 es

$$C_1 = \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

el cofactor de c_2 es

$$C_2 = -\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}$$

y

$$C_3 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

es cofactor de c_3. ▷

EJEMPLO. Calcular el determinante

$$\begin{vmatrix} 1 & 0 & -1 \\ 4 & 0 & -1 \\ 0 & 2 & 4 \end{vmatrix}$$



desarrollando por la segunda columna.

◁ Utilizando las paridades y el teorema del desarrollo 1.3 se tiene que

$$\begin{vmatrix} 1 & 0 & -1 \\ 4 & 0 & -1 \\ 0 & 2 & 4 \end{vmatrix} = -0\begin{vmatrix} 4 & -1 \\ 0 & 4 \end{vmatrix} + 0\begin{vmatrix} 1 & -1 \\ 0 & 4 \end{vmatrix} - 2\begin{vmatrix} 1 & -1 \\ 4 & -1 \end{vmatrix}$$

$$= -2(-1+4) = -6. \quad \triangleright$$

EJEMPLO. Calculamos el conocido determinante

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$$

desarrollando por el segundo renglón.

$$\triangleleft \quad \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix} = -2\begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} + 3\begin{vmatrix} 1 & 3 \\ 3 & 5 \end{vmatrix} - 4\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$$

$$= -2(10-12) + 3(5-9) - 4(4-6) = 4 - 12 + 8 = 0. \quad \triangleright$$

LEMA 1.2 *Si un determinante cambia los renglones por columnas, no se altera tal. Esto es, el determinante de una matriz y el determinante de su transpuesta coinciden.*

◁ Por el teorema del desarrollo tenemos

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1A_1 + b_1B_1 + c_1C_1 = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \quad \triangleright$$

LEMA 1.3 *Si en un determinante se intercambian dos renglones (o dos columnas) el signo del determinante se cambia.*

◁ Por el teorema del desarrollo se tiene

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1A_1 + b_1B_1 + c_1C_1$$

$$= a_1(b_2c_3 - c_2b_3) - b_1(a_2c_3 - a_3c_2) + c_1(a_2b_2 - a_3b_2)$$

Por otro lado, al calcular el determinante después de intercambiar los dos primeros renglones, obtenemos

$$\begin{vmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_2(b_1c_3 - c_1b_3) - b_2(a_1c_3 - a_3c_1) + c_2(a_1b_3 - a_3b_1)$$

$$= -a_1(b_2c_3 - c_2b_3) + b_1(a_2c_3 - c_2a_3) - c_1(a_2b_3 - a_3b_2)$$

$$= - \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \quad \triangleright$$

EJEMPLO. Ya hemos calculado el determinante

$$\begin{vmatrix} 1 & 0 & -1 \\ 4 & 0 & -1 \\ 0 & 2 & 4 \end{vmatrix} = -6$$

$\triangleleft$ Si intercambiamos las columnas primera y segunda, obtenemos, al desarrollar por la nueva primer columna

$$\begin{vmatrix} 0 & 1 & -1 \\ 0 & 4 & -1 \\ 2 & 0 & 4 \end{vmatrix} = 2 \begin{vmatrix} 1 & -1 \\ 4 & -1 \end{vmatrix} = 2(-1+4) = 6 \quad \triangleright$$

Del Lema 1.3, para el determinante con dos renglones iguales, al intercambiarlos, obtenemos

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix} = - \begin{vmatrix} a_1 & b_1 & c_1 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

lo que implica que tal determinante es nulo debido a que 0 es el único número real que es igual a su negativo. Esto es,

COROLARIO 1.2 *Si en un determinante se repiten dos renglones o dos columnas el determinante vale cero.*

EJEMPLO. Por un cálculo directo, se tiene que el determinante siguiente con dos renglones iguales

$$\begin{vmatrix} 1 & 0 & 1 \\ 1 & 3 & 4 \\ 1 & 3 & 4 \end{vmatrix} = 3 \begin{vmatrix} 1 & 1 \\ 1 & 4 \end{vmatrix} - 3 \begin{vmatrix} 1 & 1 \\ 1 & 4 \end{vmatrix} = 3(4-1) - 3(4-1) = 0$$

cuando se desarrolla por la segunda columna.

Tenemos además el siguiente importante resultado sobre desarrollos.

COROLARIO 1.3 *La suma de los productos pares de los elementos de una fila o renglón por los complementos algebráicos correspondientes es cero.*

Es decir, para cualquier determinante se cumplen las igualdades

$$a_1A_2 + b_1B_2 + c_1C_2 = 0$$

$$b_1A_1 + b_2A_2 + b_3A_3 = 0$$

$$c_1A_1 + c_2A_2 + c_3A_3 = 0$$

◁ Consideremos un determinante arbitrario

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

entonces

$$A_2 = -\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix}, B_2 = \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix}, C_2 = -\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}$$

Por el colorario 1.2 el siguiente determinante se anula

$$0 = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

al desarrollar tal determinante por el segundo reglón se tiene

$$0 = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix} = -a_1\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + b_1\begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} - c_1\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}$$

$$= a_1A_2 + b_1B_2 + c_1C_2$$

lo que prueba la primer igualdad.

Las otras se prueban de forma análoga. ▷

OBSERVACIÓN. Para cualquier escalar $\lambda \in \mathbb{R}$ se tiene que

$$\lhd \quad \begin{vmatrix} \lambda a_1 & \lambda b_1 & \lambda c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \lambda a_1A_1 + \lambda b_1B_1 + \lambda c_1 = \lambda(a_1A_1 + b_1B_1 + c_1C_1)$$

$$= \lambda\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \quad \rhd$$

Esto se resume en el siguiente resultado.

LEMA 1.4 *Se puede sacar fuera del determinante un factor común de los elementos de un renglón (o una columna).*

OBSERVACIÓN. Utilizando el Lema 1.4 con $\lambda = 0$, obtenemos,

$$\lhd \quad \begin{vmatrix} 0 & 0 & 0 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} 0a_1 & 0b_1 & 0c_1 \\ a_2 & b_2 & c_3 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0 \begin{vmatrix} a_1 & b_1 & c_1 \\ a_3 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0 \quad \rhd$$

lo que nos conduce al siguiente corolario.

COROLARIO 1.4 *Si en un determinante un renglón (o una columna) tiene todos sus elementos cero, entonces su valor es cero.*

OBSERVACIÓN. Sea λ un número real tal que $a_1 = \lambda a_2, b_1 = \lambda b_2, c_1 = \lambda c_2$, entonces de acuerdo al Lema 1.4 y al corolario 1.4,

$$\lhd \quad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} \lambda a_2 & \lambda b_2 & \lambda c_2 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \lambda \begin{vmatrix} a_2 & b_2 & c_2 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0 \quad \rhd$$

Esto implica el siguiente corolario.

COROLARIO 1.5 *Si en un determinante un renglón (o una columna) es proporcional a otro renglón (o a otra columna) entonces tal vale cero.*

EJEMPLO. Al calcular el determinante

$$\begin{vmatrix} 1 & 0 & 1 \\ 1 & 3 & 4 \\ 4 & 12 & 16 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 \\ 1 & 3 & 4 \\ 4(1) & 4(3) & 4(4) \end{vmatrix} = 4 \begin{vmatrix} 1 & 0 & 1 \\ 1 & 3 & 4 \\ 1 & 3 & 4 \end{vmatrix} = 4(0) = 0$$

en virtud que el último determinante tiene dos renglones repetidos y se anula.

OBSERVACIÓN. Sean $\lambda_1, \beta_1, \gamma_1$ números reales arbitrarios, entonces

$$\lhd \quad \begin{vmatrix} a_1 + \lambda_1 & b_1 + \beta_1 & c_1 + \gamma_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = (a_1+\lambda_1)A_1 + (b_1+\beta_1)B_1 + (c_1+\gamma_1)C_1$$

$$= (a_1A_1 + b_1B_1 + c_1C_1) + (\lambda_1A_1 + \beta_1B_1 + \gamma_1C_1)$$

$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} \lambda_1 & \beta_1 & \gamma_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \quad \rhd$$

Resumimos la discusión previa en el siguiente:

COROLARIO 1.6 *Si en un determinante un renglón (o una columna) los elementos son sumas, entonces tal determinante puede ser descompuesto como la suma de los determinantes de los sumandos correspondientes (respetando factores).*

EJEMPLO. Calcular el determinante

$$\begin{vmatrix} 4 & 5 & 6 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{vmatrix}$$

$\lhd$ Claramente se puede descomponer en la siguiente cadena.

$$\begin{vmatrix} 4 & 5 & 6 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 2+2 & 2+3 & 3+3 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{vmatrix} + \begin{vmatrix} 2 & 3 & 3 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix}$$

$$= \left(2\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} - \begin{vmatrix} 2 & 3 \\ 1 & 0 \end{vmatrix} - 1\begin{vmatrix} 2 & 3 \\ 1 & 0 \end{vmatrix}\right) + \left(-3\begin{vmatrix} 1 & 1 \\ -1 & 0 \end{vmatrix} - \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix}\right)$$

$$= (2(-1) + 3 - 2) + (3 - (2 - 3)) = (-1) + 4 = 3$$

donde los determinantes en la última suma se han calculado, el primero desarrollando por la primer columna, y el segundo por su segunda columna. $\rhd$

COROLARIO 1.7 *El valor del determinante no cambia cuando a un renglón (o una columna) se le agregan proporcionales de otro renglón (columna) paralelo.*

$\lhd$ Sea $\lambda \in \mathbb{R}$ un escalar arbitrario, entonces, ya que un determinante con un renglón repetido es cero, tenemos que

$$\lhd \quad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \pm \lambda \begin{vmatrix} a_2 & b_2 & c_2 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} =$$

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} \pm\lambda a_2 & \pm\lambda b_2 & \pm\lambda c_2 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} =$$

$$= \begin{vmatrix} a_1 \pm \lambda a_2 & b_1 \pm \lambda b_2 & c_1 \pm \lambda c_2 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \rhd$$



Todas los resultados enunciados se llaman **transformaciones elementales** de los determinantes.

EJEMPLO. Calcular el determinante

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{vmatrix}$$

◁ Con arreglo al Colorario 1.7 esto se puede calcular por la cadena de igualdades

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 2-2(1) & 1-2(2) & 2-2(3) \\ 3-3(1) & 2-3(2) & 1-3(3) \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -4 \\ 0 & -4 & -8 \end{vmatrix} = 1\begin{vmatrix} -3 & -4 \\ -4 & -8 \end{vmatrix} = 24 - 16 = 8 \quad \triangleright$$

Estos resultados sobre transformaciones elementales de los determinantes se pueden generalizar para **órdenes** (dimensiones) más grandes. Aquí hemos detallado la teoría apenas para órdenes 2 y 3.

Tenemos el siguiente resultado para el cálculo del del determinante de un producto de matrices cuadradas.

TEOREMA 1.4 (determinante de un producto) *Sean* $A = A_{3\times 3}$, $B = B_{3\times 3}$ *dos matrices cuadradas de* 3×3, *entonces*

$$|AB| = |A||B|$$

Esto es, el determinante de un producto es el producto de los determinantes

◁ Es un cálculo directo y omitimos su demostración ▷

EJEMPLO. Sean las matrices

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

verificar la igualdad $|AB| = |A||B|$.

◁ En el anterior ejemplo se calcularon los determinantes

$$|A| = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix} = -18, \quad |B| = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 3 & 1 \\ 0 & 1 & 2 \end{vmatrix} = 5$$



de donde $|A||B| = (-18)(5) = -90$

Por otro lado,

$$AB = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 11 & 11 \\ 2 & 14 & 11 \\ 3 & 11 & 14 \end{pmatrix}$$

cuyo determinante se calcula

$$|AB| = \begin{vmatrix} 1 & 11 & 11 \\ 2 & 14 & 11 \\ 3 & 11 & 14 \end{vmatrix} = \begin{vmatrix} 1 & 11 & 11 \\ 2-2(1) & 14-2(11) & 11-2(11) \\ 3-3(1) & 11-3(11) & 14-3(11) \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 11 & 11 \\ 0 & -8 & -11 \\ 0 & -22 & -19 \end{vmatrix} = \begin{vmatrix} -8 & -11 \\ -22 & -19 \end{vmatrix} = 8(19) - 11(22) = 152 - 242 = -90$$

que prueba la igualdad mencionada. $\triangleright$

1.5 Inversa de una matriz

Damos una forma de calcular la inversa de una matriz cuadrada de 2×2, así como una condición necesaria y suficiente de su existencia.

Una matriz cuadrada

$$A = A_{2\times 2} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

es **invertible** si existe una matriz

$$B = B_{2\times 2} = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

tal que $I = AB = BA$, donde

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

es la matriz identidad de 2×2.

Se identifica $B = A^{-1}$ y se llama **matriz inversa** de A.

TEOREMA 1.5 *Si $A = A_{2\times 2}$ es tal que $|A| \neq 0$, entonces A es invertible.*



◁ Buscamos una matriz de 2×2

$$B = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

tal que $I = BA$, es decir, buscamos números reales x, y, z, w tales que

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ax + cy & bx + dy \\ az + cw & bz + dw \end{pmatrix}$$

lo que nos lleva al sistema de ecuaciones lineales

$$\begin{cases} ax + cy = 1 \\ bx + dy = 0 \\ az + cw = 0 \\ bz + dw = 1 \end{cases}$$

Primeramente resolvemos , mediante el Lema 1.1, el sistema de 2×2

$$\begin{cases} ax + cy = 1 \\ bx + dy = 0 \end{cases}$$

donde tenemos por hipótesis que $P = ad - bc \neq 0$.

Al calcular los otros determinantes se obtiene

$$D_x = \begin{vmatrix} 1 & c \\ 0 & d \end{vmatrix} = d$$

$$D_y = \begin{vmatrix} a & 1 \\ b & 0 \end{vmatrix} = -b$$

lo que implica que

$$x = \frac{D_x}{D} = \frac{d}{ad - bc}, \; y = \frac{D_y}{D} = \frac{-b}{ad - bc}$$

En forma análoga, se obtienen las otras variables

$$z = \frac{c}{bc - ad} = \frac{-c}{ad - bc}, \; w = \frac{-a}{bc - ad} = \frac{a}{ad - bc}$$

De esta manera, A^{-1} la matriz inversa de A tiene coeficientes,

$$A^{-1} = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \quad \triangleright$$

EJEMPLO. Sea la matriz

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$



◁ Calculamos el determinante de A, obteniendo

$$|A| = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 2 - 1 = 1 \neq 0$$

y por el teorema 1.5, A es invertible, y su inversa A^{-1}se calcula

$$A^{-1} = \frac{1}{1}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

Verificamos esta afirmación realizando el producto de las matrices

$$A\,A^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \quad \triangleright$$

1.6 La regla de Cramer

En esta parte generalizamos el resultado sobre la solución de un sistema de ecuaciones lineales obtenido en el Lema 1.1, para sistemas lineales de tres ecuaciones con tres incógnitas.

Consideremos ahora el sistema de tres ecuaciones lineales con coeficientes reales con tres incógnitas x, y, z

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_3y + c_3z = d_3 \end{cases}$$

◁ Introducimos, de igual forma que para los sistemas lineales de 2×2 los determinantes

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1A_1 + b_1B_1 + c_1C_1$$

$$D_x = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix} = d_1A_1 + d_2A_2 + d_3A_3$$

$$D_y = \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}$$

$$D_z = \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}$$



Si multiplicamos el primer renglón del sistema por A_1, el segundo renglón del sistema por A_2, el tercer renglón del sistema por A_3, obtenemos que

$$a_1A_1x + b_1A_1y + c_1A_1z = d_1A_1$$

$$a_2A_2x + b_2A_2y + c_2A_2z = d_2A_2$$

$$a_3A_3x + b_3A_3y + c_3A_3z = d_3A_3$$

Así, al sumar los miembros derechos e izquierdos respectivos se obtiene

$$(a_1A_1 + a_2A_2 + a_3A_3)x + (b_1A_1 + b_2A_2 + b_3A_3)y + (c_1A_1 + c_2A_2 + c_3A_3)z$$

$$= d_1A_1 + d_2A_2 + d_3A_3$$

El colorario 1.3 implica que la anterior igualdad se escribe como

$$Dx + 0y + 0z = D_x$$

De donde, si $D \neq 0$, entonces, al despejar x, se consigue el valor

$$x = \frac{D_x}{D}$$

De igual forma, podemos obtener y, z mediante un procedimiento similar (suponiendo que $D \neq 0$).

$$y = \frac{D_y}{D}, \quad z = \frac{D_z}{D}$$

Resumimos la anterior discusión en el siguiente resultado.

TEOREMA 1.6 (Regla de Cramer) *Para el sistema lineal de tres ecuaciones con tres incógnitas dado, se tiene una tripleta de soluciones dadas por las igualdades*

$$x = \frac{D_x}{D}, \quad y = \frac{D_y}{D}, \quad z = \frac{D_z}{D}$$

suponiendo que $D \neq 0$.

EJEMPLO. Resolver, utilizando La regla de Cramer, el sistema de ecuaciones

$$\begin{cases} x + 2y + 3z = 1 \\ 2x + 3y + z = 0 \\ 3x + y + 2y = 0 \end{cases}$$

◁ Calculamos los determinantes respectivos.

$$D = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 2-2(1) & 3-2(2) & 1-2(3) \\ 3-3(1) & 1-3(2) & 2-3(3) \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & -5 & -7 \end{vmatrix}$$

$$= \begin{vmatrix} -1 & -5 \\ -5 & -7 \end{vmatrix} = 7 - 25 = -18 \neq 0$$

$$D_x = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 3 & 1 \\ 0 & 1 & 2 \end{vmatrix} = 1 \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} = 6 - 1 = 5$$

$$D_y = \begin{vmatrix} 1 & 1 & 3 \\ 2 & 0 & 1 \\ 3 & 0 & 2 \end{vmatrix} = -1 \begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} = -1$$

$$D_z = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 0 \\ 3 & 1 & 0 \end{vmatrix} = 1 \begin{vmatrix} 2 & 3 \\ 3 & 1 \end{vmatrix} = -7$$

Por La regla de Cramer se obtiene finalmente que

$$x = \frac{D_x}{D} = \frac{5}{-18} \quad y = \frac{D_y}{D} = \frac{-1}{-18} = \frac{1}{18} \quad z = \frac{D_z}{D} = \frac{-7}{-18} = \frac{7}{18} \quad \triangleright$$

A continuación, iniciamos el estudio de aquellos sistemas de ecuaciones de 3×3 (tres ecuaciones por tres incógnitas) donde el determinante del sistema se anula, pero se obtienen soluciones imponiendo una condición adicional.

Consideremos el mismo sistema inicial

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_3y + c_3z = d_3 \end{cases}$$

pero, supóngase que en este caso el determinante principal se anula, esto es,

$$\widetilde{D} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0$$

Supóngase, además, que el determinante tiene el menor de 2×2

$$D = C_3 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$$



Consideremos entonces el sistema de 2×3 dado por las dos ecuaciones primeras del sistema

$$(R): \begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases}$$

Tomamos a la variable z como un parámetro t, es decir, hacemos $z = t$ y resolvemos el sistema de 2×2

$$\begin{cases} a_1x + b_1y = d_1 - c_1t \\ a_2x + b_2y = d_2 - c_2t \end{cases}$$

mediante El método de Cramer, en virtud que el determinante principal del sistema $D = a_1b_2 - a_2b_1$ no se anula, por hipótesis.

Calculemos los otros determinantes,

$$D_x = \begin{vmatrix} d_1 - c_1t & b_1 \\ d_2 - c_2t & b_2 \end{vmatrix} = \begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix} + \begin{vmatrix} -c_1 & b_1 \\ -c_2 & b_2 \end{vmatrix} t$$

$$= \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} t + \begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix} = A_3t + \begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix}$$

$$D_y = \begin{vmatrix} a_1 & d_1 - c_1t \\ a_2 & d_2 - c_2t \end{vmatrix} = \begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix} + \begin{vmatrix} a_1 & -c_1 \\ a_2 & -c_2 \end{vmatrix} t-$$

$$= \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} t + \begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix} = B_3t + \begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix}$$

de donde, por El método de Cramer, las soluciones de tal sistema (R) vienen dadas por

$$\begin{cases} x = \frac{D_x}{D} = \dfrac{A_3t + \begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix}}{D} = \dfrac{A_3t + \begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix}}{C_3} \\ y = \frac{D_y}{D} = \dfrac{B_3t + \begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix}}{C_3} \\ z = t \end{cases}$$

Afirmamos que estas soluciones resuelven la tercera ecuación también cuando el determinante

$$\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix} = 0.$$

En efecto, al sustituir los valores de x, y y z, se tiene,

$$a_3x + b_3y + c_3z = a_3\left(\frac{A_3t + \begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix}}{C_3}\right) + b_3\left(\frac{B_3t + \begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix}}{C_3}\right) + c_3t$$

$$= \frac{a_3A_3t + b_3B_3t + c_3C_3t}{C_3} + \frac{a_3\begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix} + b_3\begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix}}{C_3}$$

$$= \frac{\widetilde{D}t}{C_3} + \frac{a_3\begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix} + b_3\begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix} - d_3\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}{C_3} + \frac{d_3\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}{C_3}$$

$$= -\frac{\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}}{C_3} + \frac{d_3C_3}{C_3} = -\frac{1}{C_3}\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix} + d_3 = d_3.$$

El hecho de que ambos determinantes

$$\widetilde{D} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \quad \text{y} \quad \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}$$

se anulan, se entiende, según el cálculo anterior, de que el sistema inicial de 3×3 es **compatible**. Para este caso, las soluciones de tal sistema vienen dadas por la tripleta mencionada.

EJEMPLO. Resolver el sistema de ecuaciones

$$\begin{cases} x + y + z = 1 \\ x - y + 3z = -2 \\ 2x + 4z = -1 \end{cases}$$

$\triangleleft$ Calculamos primeramente el determinante principal

$$\widetilde{D} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 3 \\ 2 & 0 & 4 \end{vmatrix} = 2\begin{vmatrix} 1 & 1 \\ -1 & 3 \end{vmatrix} + 4\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = 2(4) + 4(-2) = 0$$

luego verificamos que tiene el menor

$$D = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -1 - 1 = -2 \neq 0.$$

Posteriormente, para verificar la compatibilidad del sistema calculamos (si se va a utilizar $z = t$ como parámetro) el determinante

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & -2 \\ 2 & 0 & -1 \end{vmatrix} = 2\begin{vmatrix} 1 & 1 \\ -1 & -2 \end{vmatrix} - 1\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = 2(-1) - (-2) = 0.$$

Según el método sugerido, consideramos el primer par de ecuaciones (donde $D \neq 0$) y hacemos $z = t$, obteniendo el sistema

$$\begin{cases} x + y = 1 - t \\ x - y = -2 - 3t \end{cases}$$

que resolvemos por El método de Cramer.

$$D_x = \begin{vmatrix} 1-t & 1 \\ -2-3t & -1 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ -2 & -1 \end{vmatrix} - \begin{vmatrix} t & 1 \\ 3t & -1 \end{vmatrix} = 1 - (-t - 3t) = 1 + 4t$$

$$D_y = \begin{vmatrix} 1 & 1-t \\ 1 & -2-3t \end{vmatrix} = -2 - 3t - (1-t) = -2 - 3t - 1 + t = -3 - 2t$$

De donde, las soluciones al sistema inicial son

$$\begin{cases} x = \frac{D_x}{D} = \frac{1+4t}{-2} = -\frac{1}{2} - 2t \\ y = \frac{D_y}{D} = \frac{-3-2t}{-2} = \frac{3}{t} + t \\ z = t \end{cases}$$

donde $t \in \mathbb{R}$ es un número real arbitrario. ▷

Lo anterior se puede resumir en el siguiente resultado.

LEMA 1.5 *Para el sistema de ecuaciones lineales*

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_3y + c_3z = d_3 \end{cases}$$

que satisface las condiciones,

a. *El determinante principal se anula:*

$$\widetilde{D} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0$$



b. *El menor* (2×2) *de* $\widetilde{D}$,

$$C_3 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$$

c. *El determinante de compatibilidad se anula,*

$$\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix} = 0$$

se tiene un **sistema parametrizado** *de soluciones*

$$\begin{cases} x = A_3 t + \dfrac{\begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix}}{C_3} \\ y = B_3 t + \dfrac{\begin{vmatrix} a_1 & d_1 \\ a_2 & b_2 \end{vmatrix}}{C_3} \\ z = t \end{cases}$$

donde $t \in \mathbb{R}$ *es un número real arbitrario.*

Observamos que el método para resolver este tipo de ecuaciones de 3×3 consiste en ignorar la tercer ecuación y resolver paramétricamente el sistema inicial de 2×3. La tercer ecuación se resuelve automáticamente con las soluciones del sistema de 2×3, sabiendo que el determinante de compatibilidad se anula.

1.7 Sistemas lineales homogéneos

En esta sección iniciamos el estudio de los sistemas de ecuaciones lineales que no tienen términos independientes. A tales sistemas les llamaremos **homogéneos**.

El procedimiento de solución de estos sistemas sigue el patrón de solución de los sistemas lineales en el Lema 1.5.

Considérese el problema por resolver del sistema homogéneo de 2×2

$$\begin{cases} 3x + 2y = 0 \\ 12x + 8y = 0 \end{cases}$$

Es claro que cuando $x = 0$, $y = 0$ se tiene una solución del sistema, llamada la **solución trivial**.



Al tratar de utilizar La regla de Cramer (Lema 1.1) no tiene sentido tal regla, pues

$$D = \begin{vmatrix} 3 & 2 \\ 12 & 8 \end{vmatrix} = 24 - 24 = 0$$

Procedemos entonces a ignorar una de las ecuaciones y nos quedamos, con la primera, por ejemplo,

$$3x + 2y = 0$$

Al despejar y, obtenemos

$$y = \frac{-3}{2}x$$

Sea $t \in \mathbb{R}$ arbitrario y pongamos $x = t$, entonces $y = \frac{-3}{2}t$.

Afirmamos que estos valores de x y y (en términos de t) también resuelven la otra ecuación.

$\lhd$ En efecto, por una simple sustitución se tiene

$$12x + 8y = 12t + 8\left(\frac{-3}{2}t\right) = 12t - 12t = 0 \quad \rhd$$

El método utilizado en este ejemplo se puede generalizar para cualquier sistema (2×2)

$$\begin{cases} ax + by = 0 \\ cx + dy = 0 \end{cases}$$

donde $ad - bc = 0$.

Consideremos ahora el sistema homogéneo de dos ecuaciones lineales con tres incógnitas

$$\begin{cases} a_1x + b_1y + c_1z = 0 \\ a_2x + b_2y + c_2z = 0 \end{cases}$$

Una simple sustitución prueba que $x = y = z = 0$ es una solución de tal sistema (**solución trivial**). Nos interesa encontrar soluciones no triviales del sistema.

Para ello, supongamos sin pérdida de generalidad que $z \neq 0$. Entonces, del sistema

$$\begin{cases} a_1x + b_1y + c_1z = 0 \\ a_2x + b_2y + c_2z = 0 \end{cases}$$

se obtiene

$$\begin{cases} a_1x + b_1y = -c_1z \\ a_2x + b_1y = -c_2z \end{cases}$$



o equivalentemente,

$$\begin{cases} \frac{a_1x+b_1y}{z} = -c_1 \\ \frac{a_2x+b_2y}{z} = -c_2 \end{cases}$$

o bien,

$$\begin{cases} a_1\left(\frac{x}{z}\right) + b_1\left(\frac{y}{z}\right) = -c_1 \\ a_2\left(\frac{x}{z}\right) + b_2\left(\frac{y}{z}\right) = -c_2 \end{cases}$$

Poniendo $v = \frac{x}{z}$, $w = \frac{y}{z}$ se obtiene

$$\begin{cases} a_1v + b_1 w = -c_1 \\ a_2v + b_2 w = -c_2 \end{cases}$$

que se resuelve por determinantes si $D = a_1b_2 - a_2b_1 \neq 0$.

$\triangleleft$ Primeramente encontramos

$$\begin{pmatrix} a_1 & b_1 & -c_1 \\ a_2 & b_2 & -c_2 \end{pmatrix}$$

la matriz principal del sistema. Después, calculamos los determinantes asociados

$$D_v = \begin{vmatrix} -c_1 & b_1 \\ -c_2 & b_2 \end{vmatrix} = -c_1b_2 + b_1c_2,$$

$$D_w = \begin{vmatrix} a_1 & -c_1 \\ a_2 & -c_2 \end{vmatrix} = -a_1c_2 + c_1a_2$$

Ya que $v = \frac{D_v}{D}$, $w = \frac{D_w}{D}$ tenemos entonces que

$$\frac{x}{z} = \frac{D_v}{D}, \quad \frac{y}{z} = \frac{D_w}{D}$$

Por lo tanto, para que x, y, z sean soluciones del sistema, se deberá cumplir la cadena de igualdades

$$t = \frac{x}{D_v} = \frac{z}{D} = \frac{y}{D_w}$$

para $t \in \mathbb{R}$. $\triangleright$

Se obtiene el siguiente resultado.

LEMA 1.6 *Todas las soluciones del sistema lineal homogéneo se obtienen por*

$$\begin{cases} x = D_vt = A_3t \\ y = D_wt = B_3t \\ z = Dt = C_3t \end{cases}$$

donde $t \in \mathbb{R}$, *y*

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0, \quad D_v = \begin{vmatrix} -c_1 & b_1 \\ -c_2 & b_2 \end{vmatrix}, \quad D_w = \begin{vmatrix} a_1 & -c_1 \\ a_2 & -c_2 \end{vmatrix}$$



EJEMPLO. Resolver el sistema lineal homogéneo

$$\left\{\begin{array}{lll} x & -2y & +3z=0 \\ 4x & +5y & -6z=0 \end{array}\right.$$

◁ Formamos la matriz del sistema

$$\begin{pmatrix} 1 & -2 & -3 \\ 4 & 5 & 6 \end{pmatrix}$$

y calculamos a que los determinantes asociados,

$$D=\begin{vmatrix} 1 & -2 \\ 4 & 5 \end{vmatrix}=5+8=13$$

$$D_v=\begin{vmatrix} -3 & -2 \\ 6 & 5 \end{vmatrix}=-15+12=-3$$

$$D_w=\begin{vmatrix} 1 & -3 \\ 4 & 6 \end{vmatrix}=6+12=18$$

De esta manera, la solución se escribe

$$\left\{\begin{array}{l} x=-3t \\ y=18t \\ z=13t \end{array}\right. , \quad t\in\mathbb{R}$$

Para ilustrar, si $t=1$, entonces $x=-3$, $y=18$, $z=13$ es solución del sistema de ecuaciones.

Si $t=\sqrt{2}$, se tendría en este caso que la tripleta $x=-3\sqrt{2}$, $y=18\sqrt{2}$, $z=13\sqrt{2}$ resuelve tal sistema. ▷

Consideremos ahora el **sistema homogéneo** de 3×3

$$\left\{\begin{array}{l} a_1x+b_1y+c_1z=0 \\ a_2x+b_2y+c_2z=0 \\ a_3x+b_3y+c_3z=0 \end{array}\right.$$

donde claramente $x=y=z=0$ es una **solución trivial** del sistema.

Nos interesa el caso cuando tal sistema tiene soluciones no triviales. Procedemos a buscar tales soluciones no triviales mediante el método utilizado para resolver, sistemas lineales en el Lema 1.5.

◁ Habíamos obtenido las ecuaciones

$$Dx=D_x, \quad Dy=D_y \quad Dz=D_z$$

en la discusión del Método de Cramer (Teorema 1.6).



Para este caso

$$D_x = \begin{vmatrix} 0 & b_1 & c_1 \\ 0 & b_2 & c_2 \\ 0 & b_3 & c_3 \end{vmatrix} = 0, \quad D_y = 0, \quad D_z = 0$$

lo que implica que

$$Dx = 0, \quad Dy = 0, \quad Dz = 0$$

De aquí que, si al menos una de las variables x, y ó z es diferente de cero tendríamos necesariamente que $D = 0$.

Recíprocamente si $D = 0$ entonces alguna de las variables x, y ó z pueden ser diferentes de cero. ⊳

Se obtiene entonces el siguiente resultado.

TEOREMA 1.7 *Para que un sistema lineal homogéneo de* 3×3 *tenga soluciones no triviales es necesario y suficiente que* $D = 0$.

Ahora supongamos que $D = 0$ y que hay un menor 2×2 de D diferente de cero. Esto es, podemos suponer que

$$0 = D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \quad \text{y} \quad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$$

sin pérdida de generalidad.

Por el Lema 1.5 el sistema homogéneo tiene solución debido a que el determinante de compatibilidad se anula, pues $d_1 = d_2 = d_3 = 0$

Por otro lado, del Lema 1.6, se tiene que el sistema homogéneo dado por el par de ecuaciones

$$\begin{cases} a_1x + b_1y + c_1z = 0 \\ a_2x + b_2y + c_2z = 0 \end{cases}$$

tiene solución

$$x = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} t = A_3t,$$

$$y = -\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} t = B_3t$$

$$z = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} t = C_3t$$

Afirmamos que x, y, z soluciones de este sistema resuelven la tercera ecuación.



En efecto, al sustituir los valores de x, y, z se cumple

$$a_3x + b_3y + c_3z = a_3(A_3t) + b_3(B_3t) + c_3(C_3t) = \\ = (a_3A_3 + b_3B_3 + c_3C_3)t = Dt = 0$$

debido a que $D = 0$ por hipótesis.

Por lo tanto,

$$x = A_3t \quad y = B_3t \quad z = C_3t$$

para $t \in \mathbb{R}$ arbitrario resuelven el sistema lineal homogéneo de 3×3. $\triangleright$

EJEMPLO. Consideremos el sistema

$$\begin{cases} x + 2y - 3z = 0 \\ 2x - y + z = 0 \\ 3x + y - 2z = 0 \end{cases}$$

$\triangleleft$ Ya que el tercer renglón de la matriz asociada es la suma de los dos primeros tenemos que

$$\begin{vmatrix} 1 & 2 & -3 \\ 2 & -1 & 1 \\ 3 & 1 & -2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -3 \\ 2 & -1 & 1 \\ 1+2 & 2-1 & -3+1 \end{vmatrix} =$$

$$= \begin{vmatrix} 1 & 2 & -3 \\ 2 & -1 & 1 \\ 1 & 2 & -3 \end{vmatrix} + \begin{vmatrix} 1 & 2 & -3 \\ 2 & -1 & 1 \\ 2 & -1 & 1 \end{vmatrix} = 0$$

pues en cada determinante se repiten renglones.

Por otro lado, el menor

$$\begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix} = -1 - 4 = -5 \neq 0,$$

lo que nos dice que para encontrar la solución, podemos considerar el par de ecuaciones

$$\begin{cases} x + 2y - 3z = 0 \\ 2x - y + z = 0 \end{cases}$$

Utilizando el Lema 1.4 se tiene que

$$x = A_3t = \begin{vmatrix} 2 & -3 \\ -1 & 1 \end{vmatrix} t = (2-3)t = -t$$

$$y = B_3t = -\begin{vmatrix} 1 & -3 \\ 2 & 1 \end{vmatrix} t = -(1+6) = -7t$$

$$z = C_3t = \begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix} t = -5t$$



es decir, la solución del sistema viene dado por la tripleta

$$\begin{cases} x = -t \\ y = -7t \\ z = -5t \end{cases},$$

donde $t \in \mathbb{R}$ es un número arbitrario. ▷

1.8 El método de Gauss-Jordan

En esta parte damos , otra metodología para resolver sistemas de ecuaciones lineales, utilizando propiedades aritméticas de las matrices con entradas reales. Como se observará, este método no apela en ningún momento a los determinantes menores del sistema, ni al orden mismo del sistema, es decir, el número de ecuaciones puede ser arbitrario, así como el número de incógnitas que intervienen en el sistema.

EJEMPLO. Resolver el sistema de ecuaciones

$$\begin{cases} 3x + 2y + z = 5 \\ x + y - z = 0 \\ 4x - y + 5z = 3 \end{cases}$$

◁ Primeramente cambiamos el orden del sistema intercambiando las primeras ecuaciones

$$\begin{cases} x + y - z = 0 \\ 3x + 2y + z = 5 \\ 4x - y + 5z = 3 \end{cases}$$

Multiplicamos por tres la primera y la restamos de la segunda

$$\begin{array}{c} 3x + 2y + z = 5 \\ 3x + 3y - 3z = 0 \\ \hline -y + 4y = 5 \end{array}$$

obteniendo el par de ecuaciones

$$\begin{cases} x + y - z = 0 \\ -y + 4z = 5 \end{cases}$$

Después, multiplicamos por la cuatro la primera del sistema inicial y la restamos de la tercera del mismo

$$\begin{array}{c} 4x - y + 5z = 3 \\ 4x + 4y - 4z = 0 \\ \hline -5y + 9z = 3 \end{array}$$

obteniendo así el sistema

$$\begin{cases} x + y - z = 0 \\ -5y + 9z = 3 \end{cases}$$

De esta manera, hemos obtenido el sistema nuevo 3×3

$$\begin{cases} x + y - z = 0 \\ -y + 4z = 5 \\ -5y + 9z = 3 \end{cases}$$

en la cual la variable x no aparece en las ecuaciones segunda y tercera.

De éste, omitimos la primera y resolvemos el sistema 2×2

$$\begin{cases} -y + 4z = 5 \\ -5y + 9z = 3 \end{cases}$$

que tiene sólo a las variables x, y.

Si multiplicamos por cinco la primer ecuación de este sistema y la restamos de la segunda obtenemos

$$\begin{array}{r} -5y + 9z = 3 \\ -5y + 20z = 25 \\ \hline 0 - 11z = -22 \end{array}$$

De esta forma, nos queda entonces el **sistema escalonado**

$$\begin{cases} x + y - z = 0 \\ -y + 4z = 5 \\ -11z = -22 \end{cases}$$

cuyo número de variables va disminuyendo en cada ecuación.

Resolviendo la tercera ecuación final para z,

$$-11z = -22$$

se obtiene que

$$z = \frac{-22}{-11} = 2$$

Al sustituir en la segunda ecuación escalonada el valor de z, se tiene que

$$5 = -y + 4z = -y + 4(2) = y + 8$$

de donde,

$$y = 8 - 5 = 3$$



Finalmente, se sustituyen los valores $y = 3$, $z = 2$ en la ecuación primera del sistema escalonado y se obtiene,

$$0 = x + y - z = x + 3 - 2 = x + 1$$

lo que nos lleva a que $x = -1$

Es decir, $x = -1$, $y = 3$, $z = 2$ resuelven el sistema inicial. $\triangleright$

OBSERVACIÓN. Si del sistema inicial se crea la matriz principal

$$\begin{pmatrix} 3 & 2 & 1 & 5 \\ 1 & 1 & -1 & 0 \\ 4 & -1 & 5 & 3 \end{pmatrix}$$

donde la última columna está formada por los elementos independientes del sistema, tal proceso se logra también escalonando tal matriz mediante **operaciones elementales** (como en los determinantes)

$$\triangleleft \quad \begin{pmatrix} 3 & 2 & 1 & 5 \\ 1 & 1 & -1 & 0 \\ 4 & -1 & 5 & 3 \end{pmatrix} \quad R_1 \to R_2, R_2 \to R_1$$

$$\sim \begin{pmatrix} 1 & 1 & -1 & 0 \\ 3 & 2 & 1 & 5 \\ 4 & -1 & 5 & 3 \end{pmatrix} (3R_1 - R_2) \to R_2, (4R_1 - R_3) \to R_3$$

$$\sim \begin{pmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & -4 & -5 \\ 0 & 5 & -9 & -3 \end{pmatrix} (5R_2 - R_3) \to R_3$$

$$\sim \begin{pmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & -4 & -5 \\ 0 & 0 & -11 & -22 \end{pmatrix}$$

Donde las expresiones $R_1 \to R_2$, $R_2 \to R_1$ definen la operación elemental de intercambiar los renglones primero y segundo en la primera equivalencia de matrices. La operación de multiplicar por tres el primer renglón y restarle el segundo para asignarlo al segundo rengón en la segunda equivalencia de matrices, se define por $(3R_1 - R_2) \to R_2$. Análogamente, la expresión $(4R_1 - R_3) \to R_3$ para la segunda equivalencia de matrices define la operación de multiplicar por cuatro el reglón primero, restarle el tercero, y asignar el resultado a la nueva matriz equivalente como un tercer renglón. Las otras expresiones designan operaciones semejantes.

De la matriz escalonada se tiene el nuevo sistema

$$\begin{cases} x + y - z = 0 \\ y - 4z = -5 \\ -11z = -22 \end{cases}$$



que se resuelve hacia atrás como en el ejemplo anterior. ▷

EJEMPLO. Resolver el sistema lineal (5×3),

$$\begin{cases} x + 2y + 3z = 14 \\ 3x + 2y + z = 10 \\ x + y + z = 6 \\ 2x + 3y + z = 5 \\ x + y = 3 \end{cases}$$

◁ Vamos a utilizar el mismo método que en el ejemplo anterior, y para ello formamos la matriz del sistema.

$$\begin{pmatrix} 1 & 2 & 3 & 14 \\ 3 & 2 & 1 & 10 \\ 1 & 1 & 1 & 6 \\ 2 & 3 & 1 & 5 \\ 1 & 1 & 0 & 3 \end{pmatrix} R_3 \to R_2,\ R_5 \to R_3\, R_2 \to R_5$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 14 \\ 1 & 1 & 1 & 6 \\ 1 & 1 & 0 & 3 \\ 2 & 3 & 1 & 5 \\ 3 & 2 & 1 & 10 \end{pmatrix} (R_2 - R_1) \to R_2,\ (R_3 - R_1) \to R_3,$$

$$(2R_1 - R_4) \to R_4\, (3R_1 - R_5) \to R_5$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 14 \\ 0 & -1 & -2 & -8 \\ 0 & -1 & -3 & -11 \\ 0 & 1 & 5 & 23 \\ 0 & 4 & 8 & 32 \end{pmatrix} \begin{matrix} (R_2 - R_3) \to R_3,\ (R_2 + R_4) \to R_4, \\ (4R_2 + R_5) \to R_5 \end{matrix}$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 14 \\ 0 & -1 & -2 & -8 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 3 & 15 \\ 0 & 0 & 0 & 0 \end{pmatrix} (3R_3 - R_4) \to R_4$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 14 \\ 1 & -1 & -2 & -8 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Como se obtiene la ecuación (**inconsistente**) del cuarto renglón final

$$0 = 0x + 0y + 0z = 1$$



el sistema no es **compatible** y, por lo tanto, el sistema inicial de 5×3 no tiene solución. ▷

EJEMPLO. Resolver el sistema de ecuaciones lineales (5×3)

$$\begin{cases} x + 2y + 3z = 14 \\ 3x + 2y + z = 10 \\ x + y + z = 6 \\ 2x + 3y - z = 5 \\ x + y = 3 \end{cases}$$

que se obtiene del sistema del ejemplo anterior al cambiar apenas el signo de z en la cuarta ecuación.

◁ Formamos la matriz del sistema y utilizamos el método anterior para escalonar la matriz, como se muestra

$$\begin{pmatrix} 1 & 2 & 3 & 14 \\ 3 & 2 & 1 & 10 \\ 1 & 1 & 1 & 6 \\ 2 & 3 & -1 & 5 \\ 1 & 1 & 0 & 3 \end{pmatrix} R_2 \longleftrightarrow R_5,$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 14 \\ 1 & 1 & 0 & 3 \\ 1 & 1 & 1 & 6 \\ 2 & 3 & -1 & 5 \\ 3 & 2 & 1 & 10 \end{pmatrix} (R_1 - R_2) \to R_2,\ (R_1 - R_3) \to R_3,$$

$$(2R_1 - R_4) \to R_4,\ (3R_1 - R_5) \to R_5$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 14 \\ 0 & 1 & 3 & 11 \\ 0 & 1 & 2 & 8 \\ 0 & 1 & 7 & 23 \\ 0 & 4 & 8 & 32 \end{pmatrix} (R_2 - R_3) \to R_3,\ (R_2 - R_4) \to R_4,\ (4R_2 - R5) \to R_5$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 14 \\ 0 & 1 & 3 & 11 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -4 & -12 \\ 0 & 0 & 4 & 12 \end{pmatrix} (R_4 + R_5) \to R_5$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 14 \\ 0 & 1 & 3 & 11 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -4 & -12 \\ 0 & 0 & 0 & 0 \end{pmatrix} (4R_3 + R_4) \to R_4$$



$$\sim \begin{pmatrix} 1 & 2 & 3 & 14 \\ 0 & 1 & 3 & 11 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

que nos lleva al sistema escalonado de 3×3,

$$\begin{cases} x + 2y + 3z = 14 \\ y + 3z = 11 \\ z = 3 \end{cases}$$

y el cual de inmediato nos da el valor de z.

Al sustituir el valor de z en la segunda ecuación, se tiene

$$y + 3(3) = 11$$

de donde, $y = 2$.

Finalmente, si se sustituyen estos valores en la primer ecuación de éste sistema último se tiene

$$14 = x + 2(2) + 3(3) = x + 13$$

con lo que $x = 1$.

Por lo tanto, la solución del sistema inicial de 5×3 viene dado por

$$x = 1, \quad y = 2, \quad z = 3 \quad \triangleright$$

Al método mostrado en los ejemplos precedentes se le llama El **método de Gauss-Jordan** para la solución de un sistema de ecuaciones de $(m \times n)$. Tal método consiste en escalonar mediante **operaciones elementales** a una matriz para obtener una matriz equivalente que defina un sistema de ecuaciones lo más simple posible.

EJEMPLO. Resolver mediante El método de Gauss-Jordan el sistema de ecuaciones

$$\begin{cases} x + 5y + 4z + 3w = 1 \\ 2x - y + 2z - w = 0 \\ 5x + 3y + 8z + w = 1 \end{cases}$$

$\triangleleft$ La matriz del sistema se escalona mediante el siguiente proceso:

$$\begin{pmatrix} 1 & 5 & 4 & 3 & 1 \\ 2 & -1 & 2 & -1 & 0 \\ 5 & 3 & 8 & 1 & 1 \end{pmatrix} \quad (2R_1 - R_2) \to R_2, \; (5R_1 - R_3) \to R_3$$

$$\sim \begin{pmatrix} 1 & 5 & 4 & 3 & 1 \\ 2 & 11 & 6 & 7 & 11 \\ 0 & 22 & 12 & 14 & 4 \end{pmatrix} (2R_2 - R_3) \to R_3$$

$$\sim \begin{pmatrix} 1 & 5 & 4 & 3 & 1 \\ 0 & -11 & 6 & 7 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Esta última matriz define al nuevo sistema de 2×4

$$\begin{cases} x + 5y + 4z + 3w = 1 \\ -11y + 6z + 7w = 2 \end{cases}$$

Resolvemos tal sistema por El método de Cramer, al efectuar el despeje de términos en las variables x, y.

$$\begin{cases} x + 5y = 1 - (4z + 3w) \\ 11y = 2 - (6z + 7w) \end{cases}$$

Primeramente, calculamos los determinantes

$$D = \begin{vmatrix} 1 & 5 \\ 0 & 11 \end{vmatrix} = 11$$

$$D_x = \begin{vmatrix} 1 - (4z + 3w) & 5 \\ 2 - (6z + 7w) & 11 \end{vmatrix}$$

$$= 11[1 - (4z + 3w)] - 5[2 - (6z + 7w)]$$

$$= 1 - 14z + 2w$$

$$D_y = \begin{vmatrix} 1 & 1 - (4z + 3w) \\ 0 & 2 - (6z + 7w) \end{vmatrix} = 2 - 6z - 7w$$

Por lo tanto,

$$x = \frac{D_x}{D} = \frac{1 - 14z + 2w}{11} = \frac{1}{11} - \frac{14}{11}z + \frac{2}{11}w$$

$$y = \frac{D_y}{D} = \frac{2 - 6z - 7w}{11} = \frac{2}{11} - \frac{6}{11}z - \frac{7}{11}w$$

Para dar una **solución paramétrica** al sistema, utilizamos dos variables auxiliares poniendo: $w = u$, $z = v$, lo que nos lleva a que la solución **biparamétrica** se escriba

$$\begin{cases} w = u \\ z = v \\ x = \frac{1}{11} - \frac{14}{11}v + \frac{2}{11}u \\ y = \frac{2}{11} - \frac{6}{11}v - \frac{7}{11}u \end{cases} \qquad \triangleright$$

El método de Gauss-Jordan involucra el concepto de **matrices semejantes**.

DEFINICIÓN. Sea $A = (a_{ij})$ una matriz real de $(n \times m)$, por una **operación elemental** en A entendemos la suma en alguno de sus renglones (columnas) de un múltiplo de un renglón paralelo (o columna paralela), un intercambio de renglones (o de columnas), o suponiendo que ningún renglón (o columna) se multiplicó por un escalar el producto de un renglón (o columna por un escalar).

EJEMPLO. En la matriz

$$\begin{pmatrix} 1 & 5 & 4 & 3 & 1 \\ 2 & -1 & 2 & -1 & 0 \\ 5 & 3 & 8 & 1 & 1 \end{pmatrix}$$

se multiplica el primer renglón por dos y se resta al segundo renglón asignando el resultado al segundo renglón definido por $(2R_1 - R_2) \to R_2$, para obtener la matriz

$$\begin{pmatrix} 1 & 5 & 4 & 3 & 1 \\ 0 & 9 & 6 & 5 & 2 \\ 5 & 3 & 8 & 1 & 1 \end{pmatrix}$$

como en el ejemplo precedente.

Los ejemplos anteriores que muestran El método de Gauss-Jordan ilustran un buen número de operaciones elementales en las matrices.

DEFINICIÓN. Dos matrices del mismo orden $(m \times n)$ se dicen ser **semejantes** si una se obtiene de la otra por un número finito de operaciones elementales.

Cuando una matriz cuadrada $A = A_{n \times n}$ es semejante a otra matriz $B = B_{n \times n}$, suponiendo que ningún renglón (o columna) se multiplicó por un escalar, el corolario 1.7 implica el siguiente resultado.

LEMA 1.7 *Si A es semejante (como se mencionó) a B, entonces sus determinantes difieren apenas por un signo.*

$\lhd$ Esto se sigue de que pudo haber intercambio de renglones. $\rhd$

De aquí se obtiene el siguiente resultado.

COROLARIO 1.8 *Si A es semejante (como se mencionó) a B y en las operaciones elementales no hay intercambio de renglones, entonces, sus determinantes coinciden, esto es,*

$$|A| = |B|$$



EJEMPLO. Consideremos la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix}$$

◁ Después de las operaciones elementales siguientes, se tiene

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix} \quad (2R_1 - R_2) \to R_2, \quad (3R_1 \to R_3) \to R_3$$

$$\sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 4 \\ 0 & 4 & 8 \end{pmatrix}$$

Con lo que

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 3 & 4 \\ 0 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 4 & 8 \\ 3 & 2 \end{vmatrix} = 24 - 16 = 8$$

Esto ya se ha calculado en ejemplo posterior al corolario 1.7 mediante la cadena de igualdades

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 2-2(1) & 1-2(2) & 2-2(3) \\ 3-3(1) & 2-3(2) & 1-3(3) \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -4 \\ 0 & -4 & -8 \end{vmatrix}$$

$$= \begin{vmatrix} -3 & -4 \\ -4 & -8 \end{vmatrix} = \begin{vmatrix} 3 & 4 \\ 4 & 8 \end{vmatrix} = 24 - 16 = 8$$

Estas igualdades involucran (salvo signos) a las matrices semejantes, donde no se ha alterado por factor alguno a los renglones, ni se han permutado. ▷

Utilizando El método de Gauss-Jordan para escalonar una matriz, procedemos a encontrar un método del cálculo de inversa multiplicativa de una matriz cuadrada, cuando es posible.

Consideremos un sistema de ecuaciones de (3×3)

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_3y + c_3z = d_3 \end{cases}$$



Debido a la regla de multiplicación de matrices, tal sistema se puede escribir matricialmente como

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

Esto es, si definimos por

$$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad d = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

lo anterior se escribiría simplemente por la ecuación matricial

$$AX = d$$

Así, el sistema de ecuaciones inicial se transforma en la ecuación matricial mencionada.

EJEMPLO. Considere el sistema de ecuaciones lineales

$$\begin{cases} 2x + 3y - z = -1 \\ 6x - 7z + 2z = 0 \\ 4x + y - 11z = -2 \end{cases}$$

$\lhd$ Este sistema se puede escribir como la ecuación matricial

$$\begin{pmatrix} 2 & 3 & -1 \\ 6 & -7 & 2 \\ 4 & 1 & -11 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \quad \rhd$$

Consideremos entonces un sistema matricial

$$AX = d$$

asociado a un sistema de ecuaciones, donde A es una matriz cuadrada de 3×3.

Si existiera una **matriz inversa** B de A, esto es, una matriz B tal que

$$I = AB = BA$$

entonces del sistema matricial se obtiene, al multiplicar por B la izquierda de cada miembro, la relación

$$BAX = Bd$$



Esto nos lleva, en virtud que $BA = I$, a que se cumpla la igualdad

$$IX = Bd.$$

Por otro lado, la matriz identidad I satisface que

$$IX = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = X$$

lo que nos lleva finalmente a la ecuación matricial

$$X = Bd$$

que indica los valores de x, y, z que resuelven el sistema lineal dado. Esto pues, si la matriz B se escribe por

$$B = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix}$$

entonces la ecuación matricial $X = Bd$ que resuelve el sistema se escribe como

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 d_1 + \beta_1 d_2 + \gamma_1 d_3 \\ \alpha_2 d_1 + \beta_2 d_2 + \gamma_2 d_3 \\ \alpha_3 d_1 + \beta_3 d_2 + \gamma_3 d_3 \end{pmatrix}$$

o equivalentemente,

$$\begin{cases} x = \alpha_1 d_1 + \beta_1 d_2 + \gamma_1 d_3 \\ y = \alpha_2 d_1 + \beta_2 d_2 + \gamma_2 d_3 \\ z = \alpha_3 d_1 + \beta_3 d_2 + \gamma_3 d_3 \end{cases}$$

Esta metodología de solución de un sistema de (3×3) nos conduce al problema de calcular la **matriz inversa** B de la matriz principal del sistema A, en caso de que tal matriz inversa exista. Una condición de existencia de la matriz B se obtiene a través de la generalización natural del teorema 1.6 para el caso de matrices 3×3.

TEOREMA 1.8 *Si $A = A_{3\times 3}$ es tal que su determinante no se anula: $|A| \neq 0$, entonces A es invertible. Esto es, existe una matriz $B = B_{3\times 3}$ tal que $AB = I = BA$, donde I es la matriz identidad de 3×3.*



EJEMPLO. Considere la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix}$$

◁ Un cálculo directo nos muestra que

$$|A| = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 5 & 3 \end{vmatrix} + 8 \begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} = -9 + 8(1) = -1$$

lo que nos dice que A es una **matriz invertible** ▷

EJEMPLO. Considere la matriz

$$A = \begin{pmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{pmatrix}$$

◁ Al calcular su determinante mediante operaciones elementales se tiene

$$\begin{vmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 6 & 4 \\ 0 & 8 & 9 \\ 0 & 8 & 9 \end{vmatrix} = 0$$

lo que hace que A no sea una matriz invertible. ▷

Analicemos por un momento el sistema matricial asociado al sistema de ecuaciones

$$AX = d$$

y la ecuación matricial de soluciones

$$X = Bd$$

Estas podrían escribirse por la pareja

$$\begin{cases} AX = Id \\ IX = Bd \end{cases}$$

donde I es la matriz identidad de 3×3.

Un procedimiento general del Método de Gauss-Jordan para resolver el sistema nos llevaría a considerar la matriz siguiente de 3×6

$$M = \begin{pmatrix} a_1 & b_1 & c_1 & 1 & 0 & 0 \\ a_2 & b_2 & c_2 & 0 & 1 & 0 \\ a_3 & b_3 & c_3 & 0 & 0 & 1 \end{pmatrix}$$



y mediante operaciones elementales llevarla (según la segunda ecuación matricial) a la matriz de 3×6:

$$N = \begin{pmatrix} 1 & 0 & 0 & \alpha_1 & \beta_1 & \gamma_1 \\ 0 & 1 & 0 & \alpha_2 & \beta_2 & \gamma_2 \\ 0 & 0 & 1 & \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix}$$

que resolvería el sistema de ecuaciones dado.

En otras palabras, dada una matriz $A_{3\times 3}$ que se pueda invertir, el proceso de llevar a la matriz M en la matriz N mediante operaciones elementales nos permite calcular la matriz inversa B de A, cuando se considera la submatriz 3×3 de N que se conforma desde la cuarta columna:

$$B = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix}$$

Este método nos permite entonces calcular la inversa de una matriz cuadrada si es invertible. En el caso de que no sea invertible la matriz inicial, el método mismo nos lleva a un absurdo durante los cálculos de las operaciones elementales. Esto se ilustra en los siguientes ejemplos.

EJEMPLO. Calcular la inversa de la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix}$$

$\lhd$ Ya hemos calculado $|A| = -1$, lo que indica que A es invertible.

Construimos la matriz de Gauss-Jordan y procedemos a realizar operaciones elementales.

$$\begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{pmatrix} (2R_1 - R_2) \to R_2,\ (R_1 - R_3) \to R_3$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & 3 & 2 & -1 & 0 \\ 0 & 2 & -5 & 1 & 0 & -1 \end{pmatrix} (-R_2) \to R_2$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 2 & -5 & 1 & 0 & -1 \end{pmatrix} (2R_2 - R_3) \to R_3$$



$$\sim \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{pmatrix} (-R_3) \to R_3$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{pmatrix} (3R_3 + R_2) \to R_2,\ (3R_3 - R_1) \to R_1$$

$$\sim \begin{pmatrix} 1 & -2 & 0 & 14 & -6 & -3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{pmatrix} (2R_2 + R_1) \to R_1$$

$$\sim \begin{pmatrix} -1 & 0 & 0 & 40 & -16 & -9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{pmatrix} (-R_1) \to R_1$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{pmatrix}$$

Esto implica que la matriz inversa A^{-1} de matriz dada A es

$$A^{-1} = \begin{pmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix}$$

Dejamos al lector verificar que $a^{-1}A = I \quad \triangleright$

Este procedimiento sirve también para calcular inversas de matrices de 2×2, como lo ilustra el siguiente ejemplo.

EJEMPLO. Sea A la matriz de 2×2 dada por

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

$\triangleleft$ Ya hemos calculado su determinante, obteniendo $|A| = 1$, lo que le hace una matriz invertible.

Construimos la matriz de Gauss-Jordan y le aplicamos operaciones elementales para calcular A^{-1}:

$$\begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \quad R_1 \leftrightarrow R_2$$

$$\sim \begin{pmatrix} 1 & 1 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix} (2R_1 - R_2) \to R_2$$



$$\sim \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & -1 & 2 \end{pmatrix} (R_2 - R_1) \to R_1$$

$$\sim \begin{pmatrix} -1 & 0 & -1 & 1 \\ 0 & 1 & -1 & 2 \end{pmatrix} (-R_1) \to R_1$$

$$\sim \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 2 \end{pmatrix}$$

Esto nos dice que la matriz inversa de A es

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \quad \triangleright$$

EJEMPLO. Consideremos la matriz

$$A = \begin{pmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{pmatrix}$$

$\triangleleft$ Ya hemos calculado $|A| = 0$, lo que nos dice que A no es una matriz invertible.

Usemos El método de Gauss-Jordan para tratar de conseguir una inversa de A.

$$\begin{pmatrix} 1 & 6 & 4 & 1 & 0 & 0 \\ 2 & 4 & -1 & 0 & 1 & 0 \\ -1 & 2 & 5 & 0 & 0 & 1 \end{pmatrix} \quad (2R_1 - R_2) \to R_2, (R_1 + R_3) \to R_3$$

$$\sim \begin{pmatrix} 1 & 6 & 4 & 1 & 0 & 0 \\ 0 & 8 & 9 & 2 & -1 & 0 \\ 0 & 8 & 9 & 1 & 0 & 1 \end{pmatrix} \quad (R_2 - R_3) \to R_3$$

$$\sim \begin{pmatrix} 1 & 6 & 4 & 1 & 0 & 0 \\ 0 & 8 & 9 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{pmatrix}$$

Ya que el renglón tercero de esta última matriz tiene ceros en sus tres primeros lugares, no es posible conseguir en adelante una matriz de forma

$$\begin{pmatrix} 1 & \alpha_1 & \beta_1 & a_1 & b_1 & c_1 \\ 0 & 1 & \beta_2 & a_2 & b_2 & c_2 \\ 0 & 0 & 1 & a_3 & b_3 & c_3 \end{pmatrix}$$

que nos permita continuar el proceso.

De aquí que el mismo Método de Gauss-Jordan pone obstrucciones por sí mismo para calcular una inversa de A, si es que A no es invertible. $\triangleright$



Ejercicios.

1. Dadas las matrices

$$A = \begin{pmatrix} -1 & 0 & 1 & 3 \\ -2 & 1 & 0 & -1 \\ 0 & 1 & -1 & 2 \end{pmatrix}_{3\times 4}, \quad B = \begin{pmatrix} 4 & 5 & 3 & 2 & 1 \\ 1 & 3 & 4 & 0 & -1 \\ -2 & 0 & 3 & 1 & 0 \end{pmatrix}_{3\times 4}$$

Realizar las operaciones matriciales $A - B$, $B - A$, $2A, -6B, 2A - 3B, (A+B)^t, A^t - 2B^t, B^t - 2A^t, A^t + \frac{1}{2}B^t$.

2. Dadas las matrices

$$C = \begin{pmatrix} 1 & -1 & 0 \\ -2 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}_{3\times 3}, \quad \begin{pmatrix} 3 & 1 & 2 \\ 3 & -1 & 1 \end{pmatrix}_{3\times 3}, \quad E = \begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 0 & 1 \end{pmatrix}_{3\times 2}$$

a. Calcular $C^2, C^3, DC, CD^t, DE, ED, (DE)^2, (ED)^2$.
b. Si A y B son como el ejercicio 1, calcular CA, A^tC, DCA, DCB.

3. Calcular, usando el teorema 1.5, la inversa de la matriz indicada.

$$A = \begin{pmatrix} 3 & -1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 2 \\ 0 & 3 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

4. Calcular los siguientes determinantes

a. Mediante la segunda columna, $\begin{vmatrix} 2 & 1 & 1 \\ 0 & 3 & 1 \\ 4 & -1 & 1 \end{vmatrix}$

b. Mediante el tercer renglón. $\begin{vmatrix} 3 & -1 & 0 \\ 1 & 2 & -1 \\ 0 & 4 & 3 \end{vmatrix}$

c. Mediante la primera columna. $\begin{vmatrix} 0 & -1 & 1 \\ 4 & 3 & 2 \\ 1 & 6 & 3 \end{vmatrix}$

d. Mediante la primera columna. $\begin{vmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{vmatrix}$



donde φ es un ángulo arbitrario.

e. $\begin{vmatrix} 2 & 3 & 4 \\ 2 & a+3 & b+4 \\ 2 & c+3 & d+4 \end{vmatrix}$ **f.** $\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix}$ **g.** $\begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{vmatrix}$

h. $\begin{vmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{vmatrix}$

5. Un número real λ se llamará un **valor propio** de la matriz cuadrada $A_{3\times3}$(ó $A_{2\times2}$) si satisface la **ecuación característica**

$$det(A - \lambda I) = 0$$

donde I es la matriz identidad. Calcular los valores propios de las siguientes matrices

a. $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ **b.** $\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$ **c.** $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$

6. Calcular los determinantes de las matrices dadas.

a. $\begin{pmatrix} -1 & 2 & 4 \\ 2 & -1 & 0 \\ 3 & 2 & 1 \end{pmatrix}$ **b.** $\begin{pmatrix} 1 & 6 & 9 \\ -1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ **c.** $\begin{pmatrix} 4 & 0 & 0 \\ -3 & 1 & 0 \\ 8 & 14 & 1 \end{pmatrix}$

d. $\begin{pmatrix} 1-t^2 & 1 & t \\ 0 & t^3 & 1 \\ 1 & 2+t & t^2 \end{pmatrix}$ **e.** $\begin{pmatrix} e^t & 2t^2+1 & t^3 \\ e^t & 4t & 3t^2 \\ e^t & 4 & 6t \end{pmatrix}$

7. Resolver mediante El método de Cramer los sistemas,

a.

$$\begin{cases} 5x - y - z = 0 \\ x + 2y + 3z = 14 \\ 4x + 3y + 2z = 16 \end{cases}$$

b.

$$\begin{cases} x + 3y - 6z = 12 \\ 3x + 2y + 5z = -10 \\ 2x + 5y + 2z = 6 \end{cases}$$

c.

$$\begin{cases} -5x + y + z = 0 \\ x - 6y + z = 0 \\ x + y - 7z = 0 \end{cases}$$



d.

$$\begin{cases} x+y+z=0 \\ 3x+6y+5z=0 \\ x+4y+3z=0 \end{cases}$$

8. Resolver mediante El método Gauss-Jordan los sistemas

a.

$$\begin{cases} 3x_1+2x_2=4 \\ x_1-4x_2=-1 \\ 7x_1+10x_2=12 \\ 5x_1+6x_2=8 \\ 3x_1-16x_2=-5 \end{cases}$$

b.

$$\begin{cases} x_1+5x_2+4x_3=1 \\ 2x_1+10x_2+8x_3=3 \\ 3x_1+15x_2+12x_3=5 \end{cases}$$

c.

$$\begin{cases} x_1-3x_2+2x_3=-1 \\ x_1+9x_2+6x_3=3 \\ x_1+3x_2+4x_3=1 \end{cases}$$

d.

$$\begin{cases} 2x_1+x_2-x_3=5 \\ x_1-2x_2+3x_3=-3 \\ 7x_1+x_2-x_3=10 \end{cases}$$

e.

$$\begin{cases} 2x_1-x_2+3x_3+2x_5=1 \\ x_1+x_2-x_3+x_4=4 \\ x_1-x_3+2x_4=6 \\ 3x_1-x_2+x_3-x_4=0 \end{cases}$$

f.

$$\begin{cases} 3x_1-x_2+x_3+2x_5=18 \\ 2x_1-5x_2+x_4+x_5=-7 \\ x_1-x_4+2x_5=6 \\ 2x_2+x_3+x_4-x_5=10 \\ x_1+x_2-3x_3+x_4=1 \end{cases}$$

g.

$$\begin{cases} 4x_1+2x_2+3x_3=-2 \\ 2x_1+8x_2-x_3=8 \\ 9x_1+x_2+8x_3=0 \end{cases}$$



9. Mediante El método de Gauss-Jordan, calcular las matrices inversas (si las hay) de

a. $\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$ **b.** $\begin{pmatrix} 2 & -3 \\ 3 & 5 \end{pmatrix}$ **c.** $\begin{pmatrix} 4 & -3 \\ 4 & 2 \end{pmatrix}$ **d.** $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

e. $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ **f.** $\begin{pmatrix} 5 & 1 & 3 \\ 1 & 4 & 2 \\ 9 & -2 & 4 \end{pmatrix}$

g. $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 2 & 4 \end{pmatrix}$ **h.** $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$

10. Una matriz A se llamará **ortogonal** si satisface que $A^t A = I$. ¿Cuáles de las siguientes matrices son ortogonales?

a. $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ **b.** $\begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$ **c.** $\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$

d. $\begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix}$ **e.** $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & -1 & -2 \end{pmatrix}$ **f.** $\begin{pmatrix} -2 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$


Casa abierta al tiempo

Capítulo 2

Vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$

2.1 Sistemas de coordenadas en $\mathbb{R}^2$ y $\mathbb{R}^3$

Comenzamos ahora el estudio de la geometría de los espacios $\mathbb{R}^2$ y $\mathbb{R}^3$ proviéndolos de un sistema de coordenadas cartesiano.

A cada punto, p en un espacio real se le asocia una n-ada de números de una manera binívoca

$$p \longleftrightarrow (x_1, x_2, \cdots, x_n)$$

El número entero n se llamará la **dimensión** del espacio.

La asociación se llamará un **sistema de coordenadas cartesiano** del n-espacio, denominado $\mathbb{R}^n$.

El número n de coordenadas dependerá de la situación de los puntos de un espacio determinado. Esto es, para asignarle una colección de números a un punto p, debemos identificar primeramente la situación del punto y el espacio donde está contenido. Tal asociación debe ser dada de manera que a cada punto en un espacio determinado le corresponde una única colección de números $x_1, \cdots, x_n$, y viceversa. Esto es, dada una n-nada de números, existe un punto único p en tal espacio con el cual está asociado.

EJEMPLO.

Para un espacio unidimensional $n = 1$, se necesita apenas una coordenada,

$$\mathbb{R}^1 : \quad p \longleftrightarrow x_1$$

Para un espacio bidimensional $n = 2$, se necesitan dos coordenadas,

$$\mathbb{R}^2 : \quad p \longleftrightarrow (x_1, x_2)$$

Para un espacio tridimensional $n = 3$, son necesarias tres coordenadas,

$$\mathbb{R}^3 : \quad p \longleftrightarrow (x_1, x_2, x_3)$$

Para el espacio cuatridimensional $n = 4$, se necesitan cuatro coordenadas,

$$\mathbb{R}^4 : \quad p \longleftrightarrow (x_1, x_2, x_3, x_4)$$

La figura 2.1 ilustra este ejemplo para $n = 1, 2, 3$.

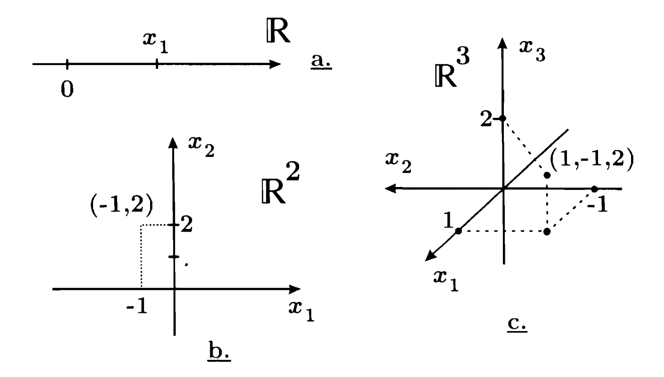


Figura 2.1: Coordenadas en $\mathbb{R}^n$ para un entero positivo n pequeño.

La localización de un punto p con sus coordenadas asociadas $x_1, \cdots, x_n$ en un espacio real $\mathbb{R}^n$ se hace mediante un sistema rectangular de ejes que tienen coincidencia en el punto común 0, y que están graduados por unidades escogidas previamente.

EJEMPLO.

a. El punto $(-1, 2)$ describe de manera única a un punto en el plano $\mathbb{R}^2$. La manera de localizarlo se muestra en la figura 2.1.

b. El punto $(1, -1, 2)$ describe una posición en el espacio $\mathbb{R}^3$. La figura 2.1 muestra la forma de localizarlo en el espacio.

c. El punto $(0, -1, 1, 4, -3)$ describe de forma única una posición en el espacio real de dimensión 5 $(\mathbb{R}^5)$.

EJEMPLO. Al estudiar un *gas ideal* se toman las variables principales P = Presión, V = Volumen, T = Temperatura.

$\triangleleft$ De esta manera, cada estado del gas se representa por una tripleta

$$(P, V, T)$$



que corresponde a un punto en $\mathbb{R}^3$. $\triangleright$

EJEMPLO. Para estudiar un sistema de cuerpos puntuales que se mueven en el espacio $\mathbb{R}^3$, es necesario determinar la posición de los cuerpos mediante un sistema de coordenadas. Así, si tenemos los cuerpos $m_1, m_2, \cdots, m_n$, a cada uno le corresponde una posición en $\mathbb{R}^3$ dada por

$$(x_1, y_1, z_1),\ (x_2, y_2, z_2), \cdots,\ (x_n, y_n, z_n)$$

respectivamente, en cada instante del movimiento. La figura 2.2 ilustra esta situación.

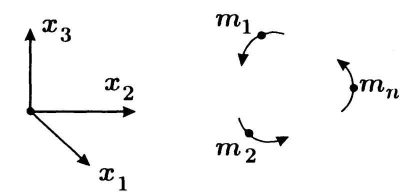


Figura 2.2: Cuerpos moviéndose en el espacio tridimensional.

A continución, le damos una **estructura algebráica** a $\mathbb{R}^n$ definiendo dos operaciones: la suma de puntos, y el producto de un número real (escalar) por un punto de $\mathbb{R}^n$.

Dados dos puntos $p = (x_1, x_2, \cdots, x_n)$, $q = (y_1, y_2, \cdots, y_n)$, en el espacio $\mathbb{R}^n$, se define el punto $p + q$ como aquél punto con coordenadas

$$p + q = (x_1 + y_1,\ x_2 + y_2, \cdots, x_n + y_n)$$

EJEMPLO. En el plano $\mathbb{R}^2$, si $p = (-1, 2)$, $q = (3, 7)$ entonces el punto $p + q \in \mathbb{R}^2$ es aquél que se obtiene mediante

$$p + q = (-1, 2) + (3, 7) = (-1 + 3,\ 2 + 7) = (2, 9)$$

EJEMPLO. En el espacio $\mathbb{R}^3$, si $p = (-1, 4, 2)$ y $q = (\pi,\ \sqrt{2},\ \frac{1}{2})$, se tiene que

$$p + q = \left(\pi - 1,\ 4 + \sqrt{2},\ \frac{5}{2}\right)$$

Dados, el punto $p = (x_1, x_2, \cdots, x_n)$ en $\mathbb{R}^n$, y el **escalar** λ, se define el nuevo punto λp como aquél en $\mathbb{R}^n$ con las coordenadas

$$\lambda p = (\lambda x_1, \cdots, \lambda x_n)$$

EJEMPLO.

a. Sean $p = \left(-1, 2, \frac{1}{3}, 4\right)$ en $\mathbb{R}^4$ y $\lambda = 3$, entonces

$$\lambda p = 3(-1, 2, \frac{1}{3}, 4) = (-3, 6, 1, 12)$$

b. Sean $p = \left(1, -1, 6, \frac{1}{5}, 8\right)$ en $\mathbb{R}^5$ y $\lambda = -5$ entonces,

$$\lambda p = (-5, 5, -30, -1, -40)$$

Las operaciones definidas de esta manera tienen propiedades heredadas de las propiedades de los números reales. Esto se observa en el siguiente,

LEMA 2.1 *Si $p_1, p_2, q_1, q_2 \in \mathbb{R}^n$ y $\lambda, \mu \in \mathbb{R}$ se cumplen las igualdades*

a.

$$(p_1 + p_2) + q_1 = p_1 + (p_2 + q_1)$$

b.

$$p_1 + q_1 = q_1 + p_1$$

c.

$$\lambda(p_1 + q_1) = \lambda p_1 + \lambda q_1$$

d.

$$(\lambda + \mu)p_1 = \lambda p_1 + \mu p_1$$

e. *Si $O = (0, \cdots, 0)$ es el punto en $\mathbb{R}^n$ con coordenadas nulas, entonces, para todo punto $p \in \mathbb{R}^n$ se cumple que*

$$p + O = p$$

f. *Si se toma $\lambda = -1$, entonces al definir por $(-1)p = -p$, se cumple*

$$p + (-p) = O$$

$\triangleleft$ Son cálculos directos realizados mediante las coordenadas.

Por ejemplo, para demostrar la propiedad **b.**, si $p_1 = (x_1, x_2, \cdots, x_n)$, y $q_1 = (y_1, y_2, \cdots, y_n)$, entonces

$$p_1 + q_1 = (x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n) = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$$

$$= (y_1 + x_1, y_2 + x_2, \cdots, y_n + x_n) = (y_1, y_2 \cdots, y_n) + (x_1, x_2, \cdots, x_n)$$

$$= q_1 + p_1$$

La propiedad **c.** se demuestra mediante la cadena de igualdades:

$$\begin{aligned}
\lambda(p_1 + q_1) &= \lambda[(x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n)] \\
&= \lambda(x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n) \\
&= (\lambda(x_1 + y_1), \lambda(x_2 + y_2), \cdots, \lambda(x_n + y_n)) \\
&= (\lambda x_1 + \lambda y_1, \lambda x_2 + \lambda y_2, \cdots, \lambda x_n + \lambda y_n) \\
&= (\lambda x_1, \lambda x_2, \cdots, \lambda x_n) + (\lambda y_1, \lambda y_2 \cdots, \lambda y_n) \\
&= \lambda(x_1, x_2, \cdots, x_n) + \lambda(y_1, y_2, \cdots, y_n) = \lambda p_1 + \lambda q_1
\end{aligned}$$

Las otras pruebas son análogas y se omiten. $\triangleright$

Vamos ahora a dar una interpretación geométrica en $\mathbb{R}^2$ de las operaciones mencionadas.

Dados los puntos $p = (x_1, y_1)$, $q = (x_2, y_2)$ estos se pueden localizar en el plano $\mathbb{R}^2$ de la forma que se mencionó anteriormente.

La suma $p + q$ se comporta como la suma de vectores ordinarios (**Ley del paralelogramo**).

Un vector es un objeto que tiene una magnitud, dirección y sentido. Se les suele describir gráficamente por flechas que tienen extremos en puntos situados en un espacio determinado.

Para sumar dos vectores según la regla del paralelogramo, se fijan en un extremo común y se desliza paralelamente uno a través del otro hasta el otro extremo, creando al realizar el deslizamiento del otro vector paralelamente, un paralelogramo.

La suma de los mencionados vectores es el vector diagonal del paralelogramo que nace en el extremo común inicial (véase la figura 2.3).

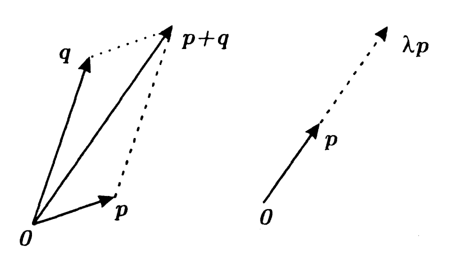


Figura 2.3: Geometría de las operaciones en $\mathbb{R}^n$.

A cada punto p y q se les asocia un **vector de posición** (nace en el origen de coordenadas $(0, 0)$) y a la suma $p + q$ se le asocia el vector

resultante de suma de vectores mediante la regla del paralelogramo (véase figura 2.3).

De igual forma, un vector en $\mathbb{R}^2$ puede ser deformado sobre su dirección cuando se multiplica por un escalar. Esto hace que varíen, su sentido o su magnitud, pero conservando su dirección.

La geometría de la operaciones mencionadas se puede generalizar a $\mathbb{R}^n$. Esto es, si $p = (x_1, x_2, \cdots, x_n)$ $q = (y_1, y_2, \cdots, y_n)$ entonces el punto

$$p + q = (x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n) = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$$

puede ser visto como el **vector resultante** de la suma por la regla del paralelogramo.

Por otro lado, si se considera el vector de posición en el punto

$$p = (x_1, x_2, \cdots, x_n)$$

entonces

$$\lambda p = \lambda(x_1, x_2, \cdots, x_n) = (\lambda x_1, \lambda x_2, \cdots, \lambda x_n)$$

tiene un vector de posición que está en la misma dirección que p, satisfaciendo

i. si $0 < \lambda < 1$, entonces λp es un encogimiento de p, pero tiene el mismo sentido.

ii. si $1 < \lambda$, entonces λp es un estiramiento de p en el mismo sentido.

iii. si $\lambda < 0$, es λp un vector en sentido inverso de p.

Por lo anterior, a los puntos asociados con vectores de posición les llamamos también **vectores**.

Así, al conjunto

$$\mathbb{R}^n = \{(x_1, x_2, \cdots, x_n) | \, x_i \in \mathbb{R}\}$$

provisto de la suma de puntos $p + q$ y el producto de un punto por un escalar λp le llamaremos el **espacio vectorial real** de dimensión n. Consecuentemente, a los puntos de $\mathbb{R}^n$ les llamaremos también **vectores**.

NOTACIÓN. En adelante se definen a los puntos de un espacio real específico por las letras

$$p, q, \cdots$$

poniendo subíndices en sus coordenadas.

Para ejemplos

$$p = (p_1, p_2, \cdots, p_n)$$



$$q = (q_1, q_2, \cdots, q_n)$$

Por otro lado, cuando se trate de especificar un punto como un vector adoptaremos la notacion por letras del tipo

$$\xi, \zeta, \eta, \cdots$$

poniendo superíndices en sus coordenadas.

Para ejemplos

$$\xi = (\xi^1, \xi^2, \cdots, \xi^n)$$

$$\zeta = (\zeta^1, \zeta^2, \cdots, \zeta^n)$$

La ventaja de reconocer a un punto de un vector mediante la notación se observa en las aplicaciones. Por ejemplo, al momento de estudiar a un vector que actúa en un punto se hace tal distinción. Esto se concreta a continuación.

Dado un punto $p \in \mathbb{R}^2$ arbitrario y otro q (móvil), se define un **vector localizado** ξ, o de **posición**, como el vector que nace en el punto p y tiene extremo final en el punto q.

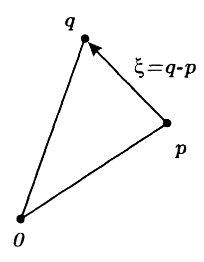


Figura 2.4: Vector localizado en un punto.

◁ Por la regla del paralelogramo se tiene de la figura 2.4, que

$$\xi + p = q$$

de donde

$$\xi = q - p$$

que nos indica que, el vector localizado ξ está definido por la diferencia $q - p$ ▷

Este concepto de vector localizado se puede generalizar para $p, q, \xi \in \mathbb{R}^n$, sin ninguna dificultad.

EJEMPLO. Dados $p = (1, 2, -1)$ y $q = (-3, 5, 0)$ en el espacio tridimensional, el vector localizado ξ que comienza en p y tiene extremo en q se calcula por

$$\lhd \quad \xi = q - p = (-3, 5, 0) - (1, 2, -1) = (-4, 3, 1) \quad \rhd$$

Por otro lado, dados un punto $p \in \mathbb{R}^n$ y ξ un vector localizado, con un extremo inicial en p, se puede conseguir siempre un punto q tal que $\xi = q - p$.

Observamos que dos vectores de posición $\xi, \eta, \in \mathbb{R}^n$ son **paralelos** si son proporcionales (están en la misma dirección), es decir, si son múltiplos por un escalar λ, es decir, $\xi = \lambda\eta$.

Por lo tanto, se tiene que dos vectores localizados son paralelos si son proporcionales entre sí.

2.2 El producto escalar y la norma en $\mathbb{R}^3$

Ahora damos un repaso a los elementos necesarios para poder definir a los objetos geométricos o físicos que pueden ser descritos por subconjuntos del plano o del espacio. Estos elementos son ya conocidos seguramente por el lector en los cursos básicos de geometría analítica, pero se repasan para no olvidarlos.

De los cursos básicos de geometría analítica, se sabe de la posibilidad de hacer mediciones de distancias entre dos puntos que están contenidos en $\mathbb{R}^3$, provisto de un sistema coordenado cartesiano.

Dados los puntos $p = (x_1, y_1, z_1)$ y $q = (x_2, y_2, z_2)$, si definimos por ℓ la distancia entre ellos, ésta se calcula por la fórmula

$$\ell^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2.$$

Al espacio $\mathbb{R}^3$ con la longitud entre dos puntos definida de esta forma, le llamaremos el **espacio Euclidiano** $\mathbb{R}^3$.

Observamos que en el caso $q = (0, 0, 0)$ se tiene que

$$\ell^2 = (x_1)^2 + (y_1)^2 + (z_1)^2$$

que es la distancia al cuadrado de p al origen.

De esta manera, si $p = (x_0, y_0, z_0)$ es un punto arbitrario de $\mathbb{R}^3$, entonces su distancia del origen, tambien llamada su **norma**, se calcula por la expresión

$$||p|| = \sqrt{x_0^2 + y_0^2 + z_0^2}$$

Una manera de inducir una longitud (norma) en $\mathbb{R}^3$ es a través de la definición de un **producto escalar.**

Esto está motivado de la relación para $p = (x, y, z)$,

$$||p||^2 = (x)^2 + (y)^2 + (z)^2 = (x)(x) + (y)(y) + (z)(z)$$

que puede definirse por

$$(x)(x) + (y)(y) + (z)(z) =< (x, y, z), (x, y, z) >=< p, p >$$

donde $< p, p >$ se define por la primer expresión en la cadena de igualdades.

Aquí hacemos la primer distinción entre los puntos y los vectores. Más que definir la operación $<\, ,\, >$ para puntos, es más conveniente definirla para vectores, con motivo de las aplicaciones.

DEFINICIÓN. Dados los vectores $\xi = (\xi^1, \xi^2, \xi^3)$, $\eta = (\eta^1, \eta^2, \eta^3)$ en el espacio cartesiano $\mathbb{R}^3$, se define el **producto escalar** de ellos, definido por $< \xi, \eta >$, como el número real (escalar)

$$< \xi, \eta >= \sum_{i=1}^{n} \xi^i \eta^i = \xi^1\eta^1 + \xi^2\eta^2 + \xi^3\eta^3.$$

Observemos que el producto escalar es una operación ejercida sobre una pareja de vectores y su valor es un número real.

EJEMPLO. Si $\xi_1 = (-1, 0, 3)$ y $\xi_2 = \left(\frac{1}{2}, \sqrt{2}, 4\right)$ en $\mathbb{R}^3$, se tiene que su producto escalar es

$$< \xi_1, \xi_2 >=< (-1, 0, 3), \left(\frac{1}{2}, \sqrt{2}, 4\right) >= (-1)\left(\frac{1}{2}\right) + 0(\sqrt{2}) + (3)(4)$$

$$= \frac{-1}{2} + 12 = \frac{23}{2}$$

EJEMPLO. Dados los vectores $\xi_1 = (4, -1)$ y $\xi_2 = (1, 4)$ se tiene que su producto escalar en el plano, está dado por

$$< \xi_1, \xi_2 >=< (4, -1), (1, 4) >= (4)(1) + (-1)(4) = 4 - 4 = 0$$

El producto escalar definirá una **norma** en virtud del siguiente Lema.
[1]

[1]En algunos textos clásicos la notación utilizada para el producto escalar es mediante un punto $\bullet$, como se indica

$$< \xi, \eta >= \xi \bullet \eta$$

que hace que también se le llame **producto punto**. Nosotros utilizaremos en este trabajo la notación $<\, ,\, >$ para efectuar las operaciones que lo requieran.

LEMA 2.2 *Dados los vectores ξ, η, ξ_1, ξ_2 en el espacio Cartesiano $\mathbb{R}^3$, y los números reales λ_1, λ_2 son justas las fórmulas*

a. $< \xi, \eta >=< \eta, \xi >$

b. $< \lambda_1 \xi_1 + \lambda_2 \xi_2, \eta >= \lambda_1 < \xi_1, \eta > + \lambda_2 < \xi_2, \eta >$

c. $< \xi, \xi > \geq 0$ *para todo ξ. De hecho, $< \xi, \xi >= 0$ sí y sólo si, $\xi = 0$.*

$\lhd$ **a.** Ya que $\xi = (\xi^1, \xi^2, \xi^3)$, $\eta = (\eta^1, \eta^2, \eta^3)$, entonces

$$< \xi, \eta >=< (\xi^1, \xi^2, \xi^3), (\eta^1, \eta^2, \eta^3) >= \xi^1\eta^1 + \xi^2\eta^2 + \xi^3\eta^3$$

$$= \eta^1\xi^1 + \eta^2\xi^2 + \eta^3\xi^3 =< (\eta^1, \eta^2, \eta^3), (\xi^1, \xi^2, \xi^3) >$$

$$=< \eta, \xi >$$

b. Si $\xi_1 = (\xi_1^1, \xi_2^2, \xi_3^3), \xi_2 = (\xi_2^1, \xi_2^2, \xi_3^3)$, entonces

$$< \lambda_1\xi_1 + \lambda_2\xi_2, \eta >=< \lambda_1(\xi_1^1, \xi_1^2, \xi_1^3) + \lambda_2(\xi_2^1, \xi_2^2, \xi_2^3), (\eta^1, \eta^2, \eta^3) >$$

$$=< (\lambda_1\xi_1^1, \lambda_1\xi_1^2, \lambda_1\xi_1^3) + (\lambda_2\xi_2^1, \lambda_2\xi_2^2, \lambda_2\xi_2^3), (\eta^1, \eta^2, \eta^3) >$$

$$=< (\lambda_1\xi_1^1 + \lambda_2\xi_2^1, \lambda_1\xi_1^2 + \lambda_2\xi_2^2, \lambda_1\xi_1^3 + \lambda_2\xi_2^3), (\eta^1, \eta^2, \eta^3) >$$

$$= (\lambda_1\xi_1^1 + \lambda_2\xi_2^1)\eta^1 + (\lambda_1\xi_1^2 + \lambda_2\xi_2^2)\eta^2 + (\lambda_1\xi_1^3 + \lambda_2\xi_2^3)\eta^3$$

$$= \lambda_1(\xi_1^1\eta^1 + \xi_1^2\eta^2 + \xi_1^3\eta^3) + \lambda_2(\xi_2^1\eta^1 + \xi_2^2\eta^2 + \xi_3^3\eta^3)$$

$$= \lambda_1 < (\xi_1^1, \xi_1^2, \xi_1^3), (\eta^1, \eta^2, \eta^3) > + \lambda_2 < (\xi_2^1, \xi_2^2, \xi_2^3), (\eta^1, \eta^2, \eta^3) >$$

$$= \lambda_1 < \xi_1, \eta > + \lambda_2 < \xi_2, \eta >$$

c. Por un cálculo directo,

$$< \xi, \xi >=< (\xi^1, \xi^2, \xi^3), (\xi^1, \xi^2, \xi^3) >= (\xi^1)^2 + (\xi^2)^2 + (\xi^3)^2 \geq 0$$

De hecho, $(\xi^1)^2 + (\xi^2)^2 + (\xi^3)^2 = 0 \iff \xi^1 = \xi^2 = \xi^3 = 0$. $\rhd$

Sea $\xi \in \mathbb{R}^3$, y considérese el conjunto de vectores en $\mathbb{R}^3$,

$$e_1 = (1, 0, 0)$$

$$e_2 = (0, 1, 0)$$

$$e_3 = (0, 0, 1)$$

$\lhd$ Si $\xi = (\xi^1, \xi^2, \xi^3)$ es un vector arbitrario, al calcular el producto escalar de ξ con cada uno de estos vectores se tiene que

$$< \xi, e_1 >=< (\xi^1, \xi^2, \xi^3), (1, 0, 0) >= \xi^1$$



$$< \xi, e_2 >=< (\xi^1, \xi^2, \xi^3), (0,1,0) >= \xi^2$$
$$< \xi_1, e_3 >= \xi^3$$

es decir, obtenemos las coordenadas de ξ. ▷

En general, si $\xi = (\xi^1, \xi^2, \cdots \xi^n)$ es un vector en $\mathbb{R}^n$, y lo definimos por

$$e_i = (\underbrace{0, 0, \cdots, 1}_{i}, 0, 0)$$

al vector con 1 en la i-ésima coordenada y con las demás coordenadas nulas, se obtiene que

$$< \xi, e_i >= \xi^i$$

lo que nos da la i-ésima coordenada de ξ.

De esta forma ξ en $\mathbb{R}^3$ se puede escribir como

$$\xi = \xi^1(1,0,0) + \xi^2(0,1,0) + \xi^3(0,0,1) = (\xi^1, \xi^2, \xi^3)$$

es decir,

$$\xi = \xi^1 e_1 + \xi^2 e_2 + \xi^3 e_3$$

donde los vectores e_i son como se definieron.

EJEMPLOS.

a. El vector $\xi_1 = (-1, 4, -2)$ se puede escribir en términos de los vectores e_1, e_2, e_3 como,

$$\xi_1 = (-1, 4, -2) = (-1, 0, 0) + (0, 4, 0) + (0, 0, -2)$$
$$= -1(1,0,0) + 4(0,1,0) - 2(0,0,1) = -1e_1 + 4e_2 - 2e_3$$

b. El vector $\xi_2 = (2, 0, 3)$ se escribe como

$$\xi_2 = 2e_1 + 0e_2 + 3e_3 = 2e_1 + 3e_3 \quad \triangleright$$

Los vectores e_1, e_2, e_3 definidos de esta forma son llamados **vectores canónicos** y satisfacen:

a. En producto escalar de un vector canónico consigo mismo es la unidad,

$$< e_1, e_1 >= 1, \; < e_2, e_2 >= 1, \; < e_3, e_3 >= 1$$

b. El producto escalar entre dos de ellos diferentes es cero,

$$< e_1, e_2 >=< (1,0,0), (0,1,0) >= 0, \; < e_1, e_3 >= 0, \; < e_2, e_3 >= 0$$

El conjunto de vectores $\{e_1, e_2, e_3\}$ se dice ser la **base canónica** de $\mathbb{R}^3$.

Dos vectores ξ, η tales que $<\xi, \eta>= 0$ se llamarán **ortogonales**. Más adelante explicaremos el por qué de esta definición.

EJEMPLO.

a. Los vectores $\xi = (-1, 1, 2)$ y $\eta = (2, 2, 0)$ son ortogonales debido a que

$$<\xi, \eta>=<(-1, 1, 2), (2, 2, 0)>= -2 + 2 + 0 = 0.$$

b. Los vectores $\xi = (1, -1, 3)$ y $\eta = (1, -1, 1)$ no son ortogonales debido a que

$$<\xi, \eta>=<(1, -1, 3), (1, -1, 1)>= 1 + 1 + 3 = 5 \neq 0$$

c. Si $\xi = (\xi^1, \xi^2)$ es un vector plano (en $\mathbb{R}^2$), es fácil construir un vector ortogonal a ξ, si se cambia el orden de las coordenadas y el signo de una de ellas: $\eta = (-\xi^2, \xi^1)$.

$$\triangleleft \quad <\xi, \eta>=<(\xi^1, \xi^2), (-\xi^2, \xi^1)>= -\xi^1\xi^2 + \xi^2\xi^1 = 0.$$

Para muestra, si $\xi = (1, 2)$, entonces $\eta = (-2, 1)$ es ortogonal a ξ. $\triangleright$

En virtud del Lema 2.3 se pueden definir los conceptos de **distancia** entre dos puntos, **norma** del vector, y **ángulo** entre dos vectores.

DEFINICIÓN. Definimos la **norma** del vector $\xi = (\xi^1, \xi^2, \xi^3)$, conocida por $||\xi||$, como

$$||\xi||^2 =<\xi, \xi>=<(\xi^1, \xi^2, \xi^3), (\xi^1, \xi^2, \xi^3)>= (\xi^1)^2 + (\xi^2)^2 + (\xi^3)^2$$

EJEMPLO. Dados los vectores $\xi_1 = (-1, 0, 3)$ y $\xi_2 = \left(\frac{1}{2}, \sqrt{2}, 4\right)$, se calculan sus normas por

$$||\xi_1|| = \sqrt{<\xi_1, \xi_1>} = \sqrt{1 + 9} = \sqrt{10}$$

$$||\xi_2|| = \sqrt{<\xi_2, \xi_2>} = \sqrt{\frac{1}{4} + 2 + 16} = \sqrt{\frac{73}{4}} = \frac{\sqrt{73}}{2}$$

No es difícil comprobar que se cumplen las siguientes propiedades de la norma.

LEMA 2.3 *Para los vectores $\xi, \eta \in \mathbb{R}^3$ y los escalares $\mu, \lambda \in \mathbb{R}$ se cumple que*

a. $||\lambda\xi|| = |\lambda| \, ||\xi||$

b. *Si $< \xi, \eta >= 0$ entonces $||\xi + \eta|| = ||\xi - \eta||$. En tal caso, los vectores ξ y η se llamarán* **ortogonales o perpendiculares**.

c. *Sı ξ y η son perpendiculares, entonces es justo El* **teorema de Pitágoras** *(véase la figura 2.5 a.)*

$$||\xi + \eta||^2 = ||\xi||^2 + ||\eta||^2$$

d. *Se cumple la igualdad de Schwarz*

$$| < \xi, \eta > | \leq ||\xi|| \, ||\eta||$$

e. *Se cumple la* **desigualdad triangular**

$$||\xi + \eta|| \leq ||\xi|| + ||\eta||$$

y la ıgualdad es válida si ξ y η son múltiplos por escalares (véase la figura 2.5 b.).

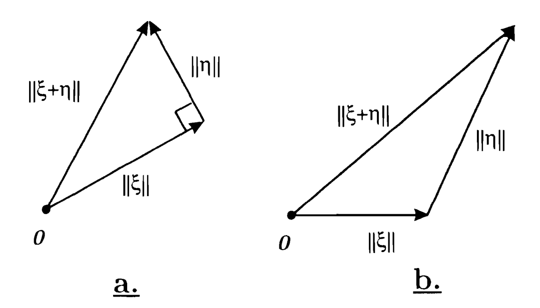


Figura 2.5: a. Teorema de Pitágoras b. Desigualdad del Triángulo.

$\triangleleft$ Demostramos algunos incisos de este Lema.

a.

$$||\lambda\xi|| = \sqrt{< \lambda\xi, \lambda\xi >} = \sqrt{\lambda^2 < \xi, \xi >} = |\lambda|\sqrt{< \xi, \xi >} = |\lambda| \, ||\xi||$$

b.

$$||\xi + \eta||^2 =< \xi + \eta, \xi + \eta >=< \xi, \xi > +2 < \xi, \eta > + < \eta, \eta >$$



$$=<\xi,\xi> + <\eta,\eta>=<\xi,\xi> -2<\xi,\eta> + <\eta,\eta>$$
$$=<\xi-\eta,\xi-\eta>= ||\xi-\eta||^2$$

c. Ya se comprobó en **b.** con las primeras tres igualdades, bajo el supuesto que $<\xi,\eta>=0$.

d. Se deja como ejercicio.

e. Por el inciso **d.**, aplicado en la segunda desigualdad siguiente

$$||\xi+\eta||^2 =<\xi,\xi> +2<\xi,\eta> + <\eta,\eta>= ||\xi||^2 + 2<\xi,\eta> +||\eta||^2$$
$$\leq ||\xi||^2 + 2|<\xi,\eta>| + ||\eta||^2 \leq ||\xi||^2 + 2||\xi||||\eta|| + ||\eta||^2$$
$$= (||\xi|| + ||\eta||)^2$$

se cumple que
$$||\xi+\eta||^2 \leq (||\xi|| + ||\eta||)^2$$
lo que implica nuestra afirmación. ▷

EJEMPLO. Sean los vectores $\xi = (-1,1,0)$, $\eta = (2,2,1)$. Entonces son perpendiculares pues $<\xi,\eta>=0$ y se verifica que

$$||\xi-\eta||^2 = ||(-3,-1,-1)||^2 = 9+1+1 = 11$$
$$||\xi+\eta||^2 = ||(1,3,1)||^2 = 1+9+1 = 11$$

es decir, se cumple el inciso a. del Lema 2.4,

$$||\xi+\eta|| = ||\xi-\eta|| = \sqrt{11}$$

Por otro lado, $||\xi||^2 = 2$, $||\eta||^2 = 9$, lo que verifica que se cumple El teorema de Pitágoras,

$$||\xi+\eta||^2 = 2+9 = ||\xi||^2 + ||\eta||^2 \quad \triangleright$$

EJEMPLO. Sean los vectores $\xi = (-1,1,1)$ y $\eta = (2,-2,1)$, entonces

$$\triangleleft \quad ||\xi+\eta|| = ||(1,-1,2)|| = \sqrt{6}$$
$$||\xi|| = \sqrt{3}, \quad ||\eta|| = \sqrt{9}$$

lo que indica que se cumple la desigualdad del triángulo para estos vectores,

$$||\xi+\eta|| = \sqrt{6} \leq \sqrt{9} \leq \sqrt{9} + \sqrt{3} = ||\eta|| + ||\xi|| \quad \triangleright$$



Un vector $\xi \in \mathbb{R}^3$ tal que $||\xi|| = 1$ se llamará **vector unitario**.

Así pues, dado el vector $\eta \in \mathbb{R}^3$ que no sea nulo, se garantiza la construcción del vector unitario

$$\xi = \frac{\eta}{||\eta||}$$

lo cual se verifica directamente,

$$\lhd \quad ||\xi|| = \left|\left|\frac{\eta}{||\eta||}\right|\right| = \frac{1}{||\eta||}||\eta|| = 1 \quad \rhd$$

EJEMPLO. Para el vector $\eta = (1, 4, -3)$ se tiene que $||\eta|| = \sqrt{26}$, con lo que se construye el vector unitario

$$\xi = \frac{(1, 4, -3)}{\sqrt{26}} = \left(\frac{1}{\sqrt{26}}, \frac{4}{\sqrt{26}}, \frac{-3}{\sqrt{26}}\right).$$

Ahora pasamos a definir la **distancia** entre los puntos arbitrarios del espacio, como la norma del vector localizado entre ellos.

Sean dados dos puntos $p = (x_1, y_1, z_1)$, $q = (x_2, y_2, z_2)$ en el espacio cartesiano, se define la distancia ℓ entre los puntos mencionados, como la norma del vector $\xi = p - q$, es decir,

$$\ell^2 = ||\xi||^2 =< p - q, p - q >$$

$$=< (x_1 - x_2, y_1 - y_2, z_1 - z_2), (x_1 - x_2, y_1 - y_2, z_1 - z_2) >$$
$$= (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2$$

EJEMPLO. Sean los puntos $p = (1, 2, 3)$ y $q = (-2, -1, 0)$ entonces la distancia entre los puntos p y q se calcula por la norma del vector $\xi = p - q = (3, 3, 3)$.

$$\lhd \quad \ell^2 = ||\xi||^2 = 3^3 + 3^2 + 3^2 = 27$$

de donde,

$$\ell = \sqrt{27} \quad \rhd$$

Sean ahora dos vectores, ξ vector unitario y η un vector arbitrario. Buscamos un vector ζ en la dirección de ξ tal que sea la **proyección ortogonal** de η en ξ, es decir, el vector $\eta - \zeta$ sea ortogonal al vector ξ como lo muestra la figura 2.6.

$\lhd$ Ya que $\eta - \zeta$ se busca ortogonal a ξ, se deberá satisfacer que $< \eta - \zeta, \xi >= 0$.

Por otro lado, si ζ está en la dirección de ξ, entonces es suficiente con encontrar un escalar λ tal que $\zeta = \lambda\xi$.

Al sustituir en la ecuación de ortogonalidad se tiene que

$$0 =< \eta - \zeta, \xi >=< \eta - \lambda\xi, \xi >=< \eta, \xi > -\lambda < \xi, \xi >$$
$$=< \eta, \xi > -\lambda||\xi||^2 =< \eta, \xi > -\lambda$$

en virtud de que ξ es unitario y por lo tanto $||\xi|| = 1$.

Esto implica que $\lambda =< \eta, \xi >=< \xi, \eta >$, y que entonces

$$\zeta =< \xi, \eta > \xi \quad \triangleright$$

Así, dados los vectores ξ unitario y η arbitrario, podemos definir la **proyección ortogonal** del vector η en el vector unitario ξ como el vector

$$\zeta =< \xi, \eta > \xi$$

que es un vector que está en la misma dirección que ξ (véase la figura 2.6).

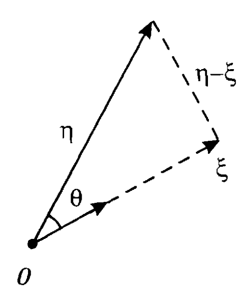


Figura 2.6: Proyección de un vector.

EJEMPLO. Realizamos la proyección del vector $\eta = (-1, 1, 1)$ en el vector $\xi = (2, -1, 5)$.

$\triangleleft$ Ya que ξ no es unitario, lo normalizamos mediante la división por su norma

$$\frac{\xi}{||\xi||} = \frac{(2, -1, 5)}{\sqrt{30}} = \left(\frac{2}{\sqrt{30}}, \frac{-1}{\sqrt{30}}, \frac{5}{\sqrt{30}}\right)$$

De esta manera, el vector buscado es

$$\zeta = \left\langle \frac{\xi}{||\xi||}, \eta \right\rangle \frac{\xi}{||\xi||} = \frac{1}{||\xi||^2} < \xi, \eta > \xi$$

$$= \frac{1}{30} < (2, -1, 5), (-1, 1, 1) > (2, -1, 5) = \frac{2}{30}(2, -1, 5) = \left(\frac{2}{15}, \frac{-1}{15}, \frac{1}{3}\right) \triangleright$$

El cálculo en este ejemplo prueba el siguiente corolario.



COROLARIO 2.1 *La* **proyección** *del vector η arbitrario en el vector no nulo ξ, llamado ζ, se calcula por la fórmula*

$$\zeta = \frac{<\xi,\eta>}{||\xi||^2}\xi$$

En muchas ocasiones, cuando no haya confusión, entenderemos no al vector ζ obtenido, sino al escalar (factor de ξ)

$$\frac{<\xi,\eta>}{||\xi||^2}$$

como la proyección mencionada.

De la figura 2.6 se tiene que si ξ,η son vectores arbitrarios y ζ es la proyección de η en ξ entonces al considerar el ángulo θ entre los vectores ξ y η se tiene que

$$\cos\theta = \frac{||\zeta||}{||\eta||}$$

Si se sustituye el valor de $\zeta = \frac{<\xi,\eta>}{||\xi||^2}\xi$ se tiene entonces que

$$\cos\theta = \frac{|<\xi,\eta>|}{||\eta||\,||\xi||^2}||\xi|| = \frac{<\xi,\eta>}{||\eta||\,||\xi||}$$

cuando consideremos a $\theta \in [0,\pi]$. Esto demuestra el siguiente Lema.

LEMA 2.4 *Si $\xi = (\xi^1,\xi^2,\xi^3)$, $\eta = (\eta^1,\eta^2,\eta^3)$ son dos vectores en $\mathbb{R}^3$, formando un ángulo θ entre ellos, se define éste angulo por la igualdad*

$$\cos\,\theta = \frac{<\xi,\,\eta>}{\sqrt{<\xi,\,\eta><\eta,\,\eta>}} = \frac{<\xi,\,\eta>}{||\xi||\,||\eta||},$$

donde $0 \leq \theta \leq \pi$.

En otras palabras el producto escalar de los vectores ξ y η se puede escribir como

$$<\xi,\eta>= ||\xi||\,||\eta||\,\cos\theta.$$

EJEMPLO. Si los vectores ξ y η cumplen que $<\xi,\eta>= 0$, sabiendo que ambos son no nulos, entonces

$$||\xi||\,||\eta||\,\cos\theta = 0$$

lo que implica que $\cos\theta = 0$, y consecuentemente $\theta = \pi/2$. Esto justifica su nombre de **vectores ortogonales**.

EJEMPLO. Sean los vectores $\xi = (2, -1, 1)$ y $\eta = (3, -4, -4)$, entonces

$$< \xi, \eta >= 6 + 4 - 4 = 6$$

y las normas son $||\xi|| = \sqrt{6}$ y $||\eta|| = \sqrt{41}$

$\triangleleft$ Por lo tanto, el ángulo formado por ξ y η es

$$\theta = \operatorname{arc\,cos} x, \left(\frac{< \xi, \eta >}{||\xi||\,||\eta||} \right) = \operatorname{arc\,cos} \frac{6}{\sqrt{6}\sqrt{41}} = \operatorname{arc\,cos} \sqrt{\frac{6}{41}}. \quad \triangleright$$

Como última observación de esta sección tenemos que de las relaciones

$$\cos\theta = \frac{< \xi, \eta >}{||\xi||||\eta||}, \quad |\cos\theta| \leq 1$$

podemos decir que

$$\left| \frac{< \xi, \eta >}{||\xi||\,||\eta||} \right| \leq 1$$

lo que nos lleva a verificar la desigualdad de Schwarz

$$| < \xi, \eta > | \leq ||\xi||\,||\eta||$$

enunciada en el inciso d. del Lema 2.3.

2.3 El producto vectorial

Definimos ahora el **producto vectorial** en el espacio $\mathbb{R}^3$.

DEFINICIÓN. Sean $\xi = (\xi^1, \xi^2, \xi^3)$, $\eta = (\eta^1, \eta^2, \eta^3)$ dos vectores en $\mathbb{R}^3$. Se define el **producto vectorial** de ξ con η definido por $[\xi, \eta]$, como el vector en $\mathbb{R}^3$ con coordenadas

$$[\xi, \eta] = (\xi^2\eta^3 - \eta^2\xi^3, -(\xi^1\eta^3 - \eta^1\xi^3), \xi^1\eta^2 - \xi^2\eta^1)$$

en la base canónica ordenada $\{e_1, e_2, e_3\}$.[2]

Una forma de recordar las coordenadas del vector $[\xi, \eta]$ es a través de la siguiente cadena convencional de igualdades

$$[\xi, \eta] = (\xi^2\eta^3 - \eta^2\xi^3, -(\xi^1\eta^3 - \eta^1\xi^3), \xi^1\eta^2 - \xi^2\eta^1) =$$

[2]En algunos textos utilizan otra notación para asegurar la operación del producto vectorial, por ejemplo: $\times$ y $\wedge$. Para ejemplificar

$$[\xi, \eta] = \xi \times \eta = \xi \wedge \eta$$

Nosotros utilizaremos la notación [,] donde se requiera efectuar la operación del producto vectorial.

$$
\begin{aligned}
&= (\xi^2\eta^3 - \eta^2\xi^3)e_1 - (\xi^1\eta^3 - \eta^1\xi^3)e_2 + (\xi^1\eta^2 - \xi^2 n^1)e_3 \\
&= e_1 \begin{vmatrix} \xi^2 & \xi^3 \\ \eta^2 & \eta^3 \end{vmatrix} - e_2 \begin{vmatrix} \xi^1 & \xi^3 \\ \eta^1 & \eta^3 \end{vmatrix} + e_3 \begin{vmatrix} \xi^1 & \xi^2 \\ \eta^1 & \eta^2 \end{vmatrix} \\
&= \begin{vmatrix} e_1 & e_2 & e_3 \\ \xi^1 & \xi^2 & \xi^3 \\ \eta^1 & \eta^2 & \eta^3 \end{vmatrix}
\end{aligned}
$$

donde el determinante último se calcula mediante la expresión penúltima.

EJEMPLO. Dados los vectores $\xi = (1, -1, 1)$ y $\eta = (-2, 3, 1)$ se tiene que

$$
[\xi, \eta] = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & -1 & 1 \\ -2 & 3 & 1 \end{vmatrix} = e_1 \begin{vmatrix} -1 & 1 \\ 3 & 1 \end{vmatrix} - e_2 \begin{vmatrix} 1 & 1 \\ -2 & 1 \end{vmatrix} + e_3 \begin{vmatrix} 1 & -1 \\ -2 & 3 \end{vmatrix}
$$

$$
= e_1(-4) - e_2(3) + e_3(1) = (-4, -3, 1)
$$

EJEMPLO. Para los vectores $\xi = (1, 1, 3)$, $\eta = (2, 2, 6)$ obtenemos el vector

$$
[\xi, \eta] = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 1 & 3 \\ 2 & 2 & 6 \end{vmatrix} = e_1 \begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} - e_2 \begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} + e_3 \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}
$$

$$
= e_1(0) - e_2(0) + e_3(0) = (0, 0, 0)
$$

En este último ejemplo no es extraño tener el resultado vector $(0, 0, 0)$, debido a que los vectores ξ y η dados son **paralelos**, es decir, ξ es un múltiplo escalar de η, de hecho, $\xi = \frac{1}{2}\eta$.

De manera general, dados dos vectores paralelos ξ y η, su producto vectorial es nulo.

$\triangleleft$ Si $\xi = \lambda\eta$ para algún $\lambda \in \mathbb{R}$, entonces,

$$
[\xi, \eta] = [\lambda\eta, \eta] = \lambda[\eta, \eta] = \lambda \begin{vmatrix} e_1 & e_2 & e_3 \\ \eta^1 & \eta^2 & \eta^3 \\ \eta^1 & \eta^2 & \eta^3 \end{vmatrix} = \lambda(0, 0, 0) = (0, 0, 0). \quad \triangleright
$$

La recíproca de esta afirmación, que toda pareja de vectores cuyo producto vectorial se anula es una pareja de vectores paralelos, es también cierta y es consecuencia del inciso f. del Lema 2.5.

Observamos que el producto vectorial $[\ ,\ ]$ está definido en parejas de vectores en $\mathbb{R}^3$ y su valor es otro vector en $\mathbb{R}^3$.

Mostramos ahora las propiedades del producto vectorial $[\ ,\ ]$ de dos vectores espaciales.

LEMA 2.5 *Dados los vectores ξ, η, ζ en $\mathbb{R}^3$ y los escalares λ, $\mu \in \mathbb{R}$ se tiene*

a.

$$[\xi, \eta] \quad \textit{es ortogonal a } \xi \textit{ y } \eta: \quad <[\xi,\eta],\xi>=0 \quad y \quad <[\xi,\eta],\eta>=0$$

b.

$$[\xi, \eta] = -[\eta, \xi], \quad \textit{(anticonmutatividad)}$$

c.

$$[\lambda\xi, \eta] = [\xi, \lambda\eta] = \lambda[\xi, \eta] \quad \textit{(conmutatividad respecto a escalar).}$$

d.

$$[\xi, \lambda\eta + \mu\zeta] = \lambda[\xi, \eta] + \mu[\xi, \zeta] \quad \textit{(distributividad)}$$

e.

$$||[\xi, \eta]||^2 = ||\xi|| ||\eta|| - (<\xi, \eta>)^2$$

f.

$$||[\xi, \eta]|| = ||\xi|| \, ||\eta|| \operatorname{sen}\theta, \quad \textit{donde } \theta \textit{ es el ángulo entre } \xi \textit{ y } \eta$$

g.

$$[\xi, [\eta, \zeta]] + [\zeta, [\xi, \eta]] + [\eta, [\zeta, \xi]] = 0 \quad \textit{(Igualdad de Jacobi)}$$

$\triangleleft$ Hacemos las pruebas apenas de algunos incisos de este Lema.

a. Si $\xi = (\xi^1, \xi^2, \xi^3)$, $\eta = (\eta^1, \eta^2, \eta^3)$ entonces

$$[\xi, \eta] = (\xi^2\eta^3 - \xi^3\eta^2, -(\xi^1\eta^3 - \xi^3\eta^1), \xi^1\eta^2 - \xi^2\eta^1)$$

Si escogemos, por ejemplo a ξ, obtenemos

$$<[\xi, \eta], \xi> = \xi^1(\xi^2\eta^3 - \xi^3\eta^2) - \xi^2(\xi^1\eta^3 - \xi^3\eta^1) + \xi^3(\xi^1\eta^2 - \xi^2\eta^1)$$
$$= \xi^1\xi^2\eta^3 - \xi^1\xi^3\eta^2 - \xi^2\xi^1\eta^3 + \xi^2\xi^3\eta^1 + \xi^3\xi^1\eta^2 - \xi^3\xi^2\eta^1 = 0$$

b. Por la propiedad de cambiar el signo a un determinante, si se permutan dos renglones se tiene

$$[\xi, \eta] = \begin{vmatrix} e_1 & e_2 & e_3 \\ \xi^1 & \xi^2 & \xi^3 \\ \eta^1 & \eta^2 & \eta^3 \end{vmatrix} = - \begin{vmatrix} e_1 & e_2 & e_3 \\ \eta^1 & \eta^2 & \eta^3 \\ \xi^1 & \xi^2 & \xi^3 \end{vmatrix} = -[\eta, \xi]$$

d. El corolario 1.7 prueba la distribución mediante la cadena de igualdades:

$$[\xi, \lambda\eta + \mu\zeta] = \begin{vmatrix} e_1 & e_2 & e_3 \\ \xi^1 & \xi^2 & \xi^3 \\ \lambda\eta^1 + \mu\zeta^1 & \lambda\eta^2 + \mu\zeta^2 & \lambda\eta^3 + \mu\zeta^3 \end{vmatrix}$$



$$= \lambda \begin{vmatrix} e_1 & e_2 & e_3 \\ \xi^1 & \xi^2 & \xi^3 \\ \eta^1 & \eta^2 & \eta^3 \end{vmatrix} + \mu \begin{vmatrix} e_1 & e_2 & e_3 \\ \xi^1 & \xi^2 & \xi^3 \\ \zeta^1 & \zeta^2 & \zeta^3 \end{vmatrix} = \lambda[\xi, \eta] + \mu[\xi, \zeta]$$

Los demás incisos se comprueban de manera similar. $\triangleright$

De la propiedad **f.** del Lema 2.5, se tiene que si dos vectores no nulos ξ, η satisfacen que su producto vectorial es el vector cero, entonces

$$0 = ||(0,0,0)|| = ||[\xi, \eta]|| = ||\xi|| \, ||\eta|| \operatorname{sen}\theta,$$

lo que implicaría que el ángulo θ formado por ellos satisface que $\operatorname{sen}\theta = 0$ o bién $\theta = \pi$. Esto es, los vectores son **paralelos**. Por esto es que podemos definir a un par de vectores como paralelos, si su producto vectorial se anula.

EJEMPLO. Sean los vectores $\xi = (1, -1, 2)$ y $\eta = (0, 1, 3)$, entonces, por el inciso **a.** del Lema 2.5 el vector $[\xi, \eta]$ es ortogonal tanto a ξ como a η. Esto es,

$$\triangleleft \quad \zeta = [\xi, \eta] = \begin{vmatrix} e_1 & e_2 & e_3 \\ 0 & -1 & 2 \\ 0 & 1 & 3 \end{vmatrix} = (-3-2)e_1 - (3)e_2 + (1)e_3 = (-5, -3, 1)$$

es ortogonal a ξ y a η. $\triangleright$

Si se consideran dos vectores arbitrarios ξ, η en $\mathbb{R}^3$, estos generan un paralelogramo R como lo muestra la figura 2.7

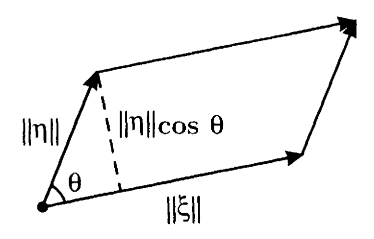


Figura 2.7: Área de un paralelogramo.

Tal paralelogramo R tiene una base $||\xi||$ y altura $||\eta|| \operatorname{sen}\theta$, donde θ es el ángulo entre ξ y η. De esta manera, el área del paralelogramo es calculado por

$$\text{Área} \quad (R) = (\text{base} \times \text{altura}) = ||\xi|| \, ||\eta|| \operatorname{sen}\theta$$

con lo que tal área se calcula según el inciso **f.** del Lema 2.5 mediante las igualdades

$$\text{Área} \quad (R) = ||\xi|| \, ||\eta|| \operatorname{sen}\theta = ||[\xi, \eta]||$$

EJEMPLO. Calcular el área del paralelogramo generado por los vectores $\xi = (7, 8, 9)$ y $\eta = (-3, 5, 2)$.

$\triangleleft$ Primeramente calculamos el vector $[\xi, \eta]$,

$$[\xi, \eta] = \begin{vmatrix} e_1 & e_2 & e_3 \\ 7 & 8 & 9 \\ -3 & 5 & 2 \end{vmatrix} = e_1 \begin{vmatrix} 8 & 9 \\ 5 & 2 \end{vmatrix} - e_2 \begin{vmatrix} 7 & 9 \\ -3 & 2 \end{vmatrix} + e_3 \begin{vmatrix} 7 & 8 \\ -3 & 5 \end{vmatrix}$$

$$= e_1(-29) - e_2(41) + e_3(59) = (-29, -41, 59).$$

y después el área del paralelogramo

$$\text{Área} = \|[\xi, \eta]\| = \sqrt{(-29)^2 + (-41)^2 + (59)^2} = \sqrt{6003} \quad \triangleright$$

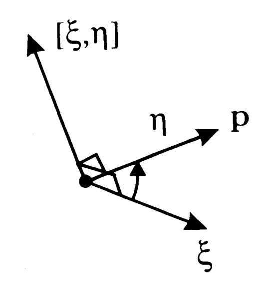


Figura 2.8: Momento de acción de ξ en η

Por otro lado, el producto vectorial puede ser considerado mecánicamente como el momento de acción del vector ξ en el punto p que tiene asociado el vector de posición η (véase la figura 2.8).

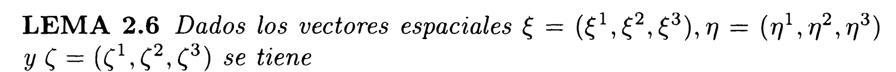
2.4 El triple producto escalar y bases de $\mathbb{R}^3$

Ahora, definimos el **producto mixto** o **triple producto escalar** de tres vectores ordenados.

DEFINICIÓN. Dados los vectores ξ, η, ζ en $\mathbb{R}^3$ se define el **producto mixto** o **triple producto escalar** de ellos al número real definido por (ξ, η, ζ) y en ese orden, como

$$(\xi, \eta, \zeta) = < [\xi, \eta], \zeta >$$

Obtenemos el siguiente resultado, útil para problemas prácticos.



LEMA 2.6 *Dados los vectores espaciales* $\xi = (\xi^1, \xi^2, \xi^3), \eta = (\eta^1, \eta^2, \eta^3)$ *y* $\zeta = (\zeta^1, \zeta^2, \zeta^3)$ *se tiene*

a. *El producto mixto de* ξ, η *y* ζ *en coordenadas se calcula por*

$$(\xi, \eta, \zeta) = \begin{vmatrix} \xi^1 & \xi^2 & \xi^3 \\ \eta^1 & \eta^2 & \eta^3 \\ \zeta^1 & \zeta^2 & \zeta^3 \end{vmatrix}$$

b. *El producto mixto no cambia si se hacen las permutaciones circulares, es decir,*

$$(\xi, \eta, \zeta) = (\zeta, \xi, \eta) = (\eta, \zeta, \xi)$$

c. *El producto mixto es anticonmutativo al transponer dos vectores, por ejemplo,*

$$(\xi, \eta, \zeta) = -(\eta, \xi, \zeta).$$

d. $(\xi, \eta, \zeta) = 0$ *si y sólo sí, los tres vectores están contenidos en un mismo plano.*

e. *El volumen* V *del paralelepípedo tendido sobre* ξ, η *y* ζ *se calcula por*

$$V = |(\xi, \eta, \zeta)|.$$

$\triangleleft$ Ya que $(\xi, \eta, \zeta) =< [\xi, \eta], \zeta >$, entonces,

$$(\xi, \eta, \zeta) = \left\langle \begin{vmatrix} e_1 & e_2 & e_3 \\ \xi^1 & \xi^2 & \xi^3 \\ \eta^1 & \eta^2 & \eta^3 \end{vmatrix}, (\zeta^1, \zeta^2, \zeta^3) \right\rangle$$

$$= \left\langle e_1 \begin{vmatrix} \xi^2 & \xi^3 \\ \eta^2 & \eta^3 \end{vmatrix} - e_2 \begin{vmatrix} \xi^1 & \xi^3 \\ \eta^1 & \eta^3 \end{vmatrix} + e_3 \begin{vmatrix} \xi^1 & \xi^2 \\ \eta^1 & \eta^2 \end{vmatrix}, (\zeta^1, \zeta^2, \zeta^3) \right\rangle$$

$$= \zeta^1 \begin{vmatrix} \xi^2 & \xi^3 \\ \eta^2 & \eta^3 \end{vmatrix} - \zeta^2 \begin{vmatrix} \xi^1 & \xi^3 \\ \eta^1 & \eta^3 \end{vmatrix} + \zeta^3 \begin{vmatrix} \xi^1 & \xi^2 \\ \eta^1 & \eta^2 \end{vmatrix}$$

$$= \begin{vmatrix} \zeta^1 & \zeta^2 & \zeta^3 \\ \xi^1 & \xi^2 & \xi^3 \\ \eta^1 & \eta^2 & \eta^3 \end{vmatrix} = \begin{vmatrix} \xi^1 & \xi^2 & \xi^3 \\ \eta^1 & \eta^2 & \eta^3 \\ \zeta^1 & \zeta^2 & \zeta^3 \end{vmatrix}$$

lo que demuestra el punto **a.**

Los puntos **b.** y **c.** tienen pruebas análogas utilizando las propiedades de los determinantes.

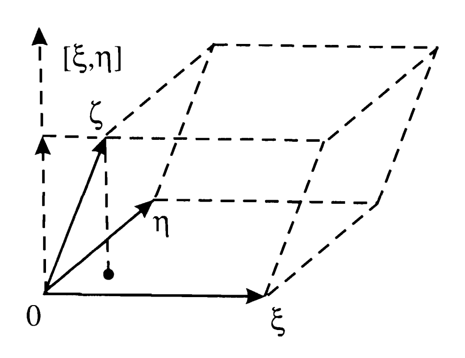


Figura 2.9: Volumen del paralelepípedo.

Para demostrar el punto **d.**, de la figura 2.9 se tiene que el paralelepípedo tiene una base con área $||[\xi,\eta]||$ y altura igual a la proyección de ζ en $[\xi,\eta]$ cuya longitud es el valor absoluto

$$\left|\frac{<[\xi,\eta],\zeta>}{||[\xi,\eta]||}\right|$$

Por lo tanto, el volumen del paralelepípedo será

$$\text{Volumen=área de la base} \times \text{altura} = ||[\xi,\eta]|| \left|\frac{<[\xi,\eta],\zeta>}{||[\xi,\eta]||}\right|$$

$$= |<[\xi,\eta],\zeta>| = |(\xi,\eta,\zeta)|$$

que prueba la afirmación.

El inciso **e.** se sigue inmediatamente del inciso **d.** ▷

EJEMPLO. Sean los vectores $\xi = (1,1,0)$, $\eta = (0,-2,3)$, $\zeta = (1,1,1)$. Entonces, el triple producto escalar de ellos se calcula

$$(\xi,\eta,\zeta) = \begin{vmatrix} 1 & 1 & 0 \\ 0 & -2 & 3 \\ 1 & 1 & 1 \end{vmatrix} = 1\begin{vmatrix} -2 & 3 \\ 1 & 1 \end{vmatrix} - 1\begin{vmatrix} 0 & 3 \\ 1 & 1 \end{vmatrix} + 0\begin{vmatrix} 0 & -2 \\ 1 & 1 \end{vmatrix}$$

$$= (-5) - (-3) + 0 = -2$$

EJEMPLO. Dados los vectores $\xi = (1,0,0), \eta = (0,2,0)$, $\zeta = (4,2,0)$ obtenemos

$$(\xi,\eta,\zeta) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 4 & 2 & 0 \end{vmatrix} = 1\begin{vmatrix} 2 & 0 \\ 2 & 0 \end{vmatrix} = 0$$



lo que era de esperarse pues ξ, η y ζ son coplanares. De hecho se observa directamente que $\zeta = 4\xi + 2\eta$.

EJEMPLO. El volumen del paralelepípedo generado por los vectores $\zeta = (2, -1, 1)$, $\eta = (1, 2, 3)$ y $\zeta = (1, 1, -2)$ es

$$V = |(\xi, \eta, \zeta)| = \left| \begin{vmatrix} 2 & -1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & -2 \end{vmatrix} \right| = |-30| = 30$$

Consideremos ahora los vectores ξ_1, ξ_2, ξ_3 en $\mathbb{R}^3$ satisfaciendo que su triple producto escalar no se anula

$$(\xi_1, \xi_2, \xi_3) \neq 0$$

Entonces, por el Lema 2.6 se tiene que no son coplanares, es decir, ninguno de ellos depende de los otros. Esto se observa del hecho de que si alguno de ellos fuera **combinación lineal** de los otros, por ejemplo

$$\xi_3 = \lambda\xi_1 + \mu\xi_2, \quad \lambda, \mu \in \mathbb{R}$$

entonces su triple producto escalar sería, mediante determinantes

$$(\xi_1, \xi_2, \xi_3) = \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{vmatrix} = \begin{vmatrix} \xi_1 \\ \xi_2 \\ \lambda\xi_1 + \mu\xi_2 \end{vmatrix} =$$

$$= \begin{vmatrix} \xi_1 \\ \xi_2 \\ \lambda\xi_1 \end{vmatrix} + \begin{vmatrix} \xi_1 \\ \xi_2 \\ \mu\xi_2 \end{vmatrix} = \lambda \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_1 \end{vmatrix} + \mu \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_2 \end{vmatrix} = 0$$

lo que no puede ser.

DEFINICIÓN. El conjunto de vectores $\{\xi_1, \xi_2, \xi_3\} \subset \mathbb{R}^3$ se dice ser **linealmente independiente** si su triple producto escalar no se anula, es decir,

$$(\xi_1, \xi_2, \xi_3) \neq 0$$

EJEMPLO. Sean los vectores canónicos $\{e_1, e_2, e_3\} \subset \mathbb{R}^3$.

$\lhd$ Es claro que

$$(e_1, e_2, e_3) = \begin{vmatrix} e_1 \\ e_2 \\ e_3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1$$

lo que los hace linealmente independientes. $\rhd$



EJEMPLO. Sea el conjunto de vectores

$$\begin{cases} \xi_1 = (1,1,0), \\ \xi_2 = (0,1,1), \\ \xi_3 = (1,1,1) \end{cases}$$

◁ Su triple producto escalar se calcula por

$$(\xi_1, \xi_2, \xi_3) = \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} = 1$$

lo que nos dice que son linealmente independientes. ▷

EJEMPLO. Sea ahora el conjunto de vectores

$$\begin{cases} \xi_1 = (1,1,0), \\ \xi_2 = (0,1,1), \\ \xi_3 = (2,5,3) \end{cases}$$

◁ El triple producto escalar entre ellos es

$$(\xi_1, \xi_2, \xi_3) == \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 5 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 5 & 3 \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ 2 & 5 \end{vmatrix} = -2 + 2 = 0$$

lo que hace de ellos un conjunto linealmente dependiente.

Esto es, existen escalares μ, λ tal que, por ejemplo,

$$\xi_3 = \lambda\xi_1 + \mu\xi_2 = \lambda(1,1,0) + \mu(0,1,1)$$

es decir,

$$(2,5,3) = (\lambda, \lambda, 0) + (0, \mu, \mu) = (\lambda, \lambda + \mu, \mu)$$

o bien,

$$\begin{cases} 2 = \lambda \\ 5 = \lambda + \mu \\ 3 = \mu \end{cases}$$

lo que implica que $\lambda = 2$, $\mu = 3$.

Por lo tanto, una **dependencia lineal** entre ellos se da mediante la igualdad

$$\xi_3 = 2\xi_1 + 3\xi_2. \quad ▷$$



EJEMPLO. Sea el conjunto $\{\xi_1(1,2), \xi_2 = (1/2, 1)\} \subset \mathbb{R}^2$, entonces

$$\begin{vmatrix} \xi_1 \\ \xi_2 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 1/2 & 1 \end{vmatrix} = 1 - 1 = 0$$

lo que hace de estos vectores un conjunto linealmente dependiente en el plano $\mathbb{R}^2$. De hecho $\xi_1 = 2\xi_2$.

EJEMPLO. El conjunto de vectores $\{\xi_1 = (1,2),\ \xi_2 = (-2,1)\}$ es linealmente independiente en $\mathbb{R}^2$ pues

$$\begin{vmatrix} \xi_1 \\ \xi_2 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ -2 & 1 \end{vmatrix} = 1 + 4 = 5 \neq 0.$$

En general, un conjunto de n-vectores en $\mathbb{R}^n$

$$\{\xi_1, \xi_2, \cdots \xi_n\}$$

se llamará **linealmente independiente** si el determinante de la matriz formada por tales vectores no se anula, es decir,

$$\begin{vmatrix} \xi_1 \\ \vdots \\ \xi_n \end{vmatrix} \neq 0$$

Sea un conjunto $\{\xi_1, \xi_2, \xi_3\}$ de vectores linealmente independientes en $\mathbb{R}^3$. Ya que no todos pertenecen a un mismo plano, al considerar cualquier vector espacial η, las proyecciones paralelas de η en cualquiera de los vectores nos provee de escalares $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tales que el vector η es una combinación lineal de ξ_1, ξ_2, ξ_2 mediante la igualdad

$$\eta = \lambda_1\xi_1 + \lambda_2\xi_2 + \lambda_3\xi_3$$

como lo muestra la figura 2.10.

Esto es, el conjunto de vectores $\{\xi_1, \xi_2, \xi_3\}$ genera al espacio $\mathbb{R}^3$ mediante combinaciones lineales.

DEFINICIÓN. El conjunto de vectores $\{\xi_1, \xi_2, \xi_3\} \subset \mathbb{R}^3$ se llamará una **base** para $\mathbb{R}^3$, si todos los vectores son linealmente independientes y generan al espacio $\mathbb{R}^3$ mediante combinaciones lineales.

Resumimos la discusión anterior en el siguiente resultado.



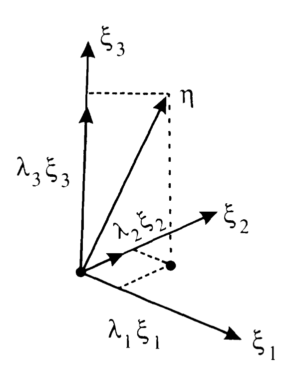


Figura 2.10: Proyección paralela de η.

TEOREMA 2.1 *Un conjunto de vectores $\{\xi_1, \xi_2, \cdots \xi_k\} \subset \mathbb{R}^n$ es una base de vectores para $\mathbb{R}^n$, si y sólo sí, $k = n$ y el determinante*

$$\begin{vmatrix} \xi_1 \\ \vdots \\ \xi_n \end{vmatrix} \neq 0$$

EJEMPLO. Los vectores canónicos espaciales

$$\{e_1 = (1,0,0),\ e_2 = (0,1,0),\ e_3 = (0,0,1)\}$$

son una base de $\mathbb{R}^3$, llamada la **base canónica**.

Cualquier vector $\eta = (\eta_1, \eta_2, \eta_3) \in \mathbb{R}^3$ se escribe

$$\eta = \eta_1 e_1 + \eta_2 e_2 + \eta_3 e_3.$$

EJEMPLO. El conjunto de vectores

$$\{\xi_1 = (1,1,0), \quad \xi_2 = (0,1,1), \quad \xi_3 = (1,1,1)\}$$

es una base de vectores para $\mathbb{R}^3$

$\triangleleft$ De hecho, si $\eta \in \mathbb{R}^3$, se tiene que existen escalares $\lambda_1, \lambda_2, \lambda_3$ tales que

$$\eta = \lambda_1\xi_1 + \lambda_2\xi_2 + \lambda_3\xi_3$$

es decir, si $\eta = (\eta^1, \eta^2, \eta^3)$, entonces

$$(\eta^1, \eta^2, \eta^3) = \lambda_1(1,1,0) + \lambda_2(0,1,1) + \lambda_3(1,1,1)$$



$$= (\lambda_1 + \lambda_3, \lambda_1 + \lambda_2 + \lambda_3, \lambda_2 + \lambda_3)$$

o equivalentemente, los escalares $\lambda_1, \lambda_2, \lambda_3$ resuelven el sistema de ecuaciones

$$\begin{cases} \lambda_1 + \lambda_3 = \eta^1 \\ \lambda_1 + \lambda_2 + \lambda_3 = \eta^2 \\ \lambda_2 + \lambda_3 = \eta^3 \end{cases}$$

Tal sistema de ecuaciones tiene determinante

$$D = \begin{vmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{vmatrix} = \left| \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}^t \right| = \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{vmatrix} = (\xi_1, \xi_2, \xi_3) \neq 0$$

lo que implica que ese sistema tiene solución para las incógnitas $\lambda_1, \lambda_2, \lambda_3$.

De hecho, por El método de Cramer

$$\lambda_1 = \frac{\begin{vmatrix} \eta^1 & 0 & 1 \\ \eta^2 & 1 & 1 \\ \eta^3 & 1 & 1 \end{vmatrix}}{D} = \frac{\begin{vmatrix} \eta^1 & \eta^2 & \eta^3 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}}{D} = \frac{(\eta, \xi_2, \xi_3)}{(\xi_1, \xi_2, \xi_3)}$$

$$\lambda_2 = \frac{\begin{vmatrix} 1 & \eta^1 & 1 \\ 1 & \eta^2 & 1 \\ 0 & \eta^3 & 1 \end{vmatrix}}{D} = \frac{\begin{vmatrix} 1 & 1 & 0 \\ \eta^1 & \eta^2 & \eta^3 \\ 1 & 1 & 1 \end{vmatrix}}{D} = \frac{(\xi_1, \eta, \xi_3)}{(\xi_1, \xi_2, \xi_3)}$$

$$\lambda_3 = \frac{(\xi_1, \xi_2, \eta)}{(\xi_1, \xi_2, \xi_3)} \quad \triangleright$$

Esto se puede generalizar en el siguiente Lema, enunciado apenas para el caso tridimensional.

LEMA 2.7 *Si el conjunto $\{\xi_1, \xi_2, \xi_3\} \subset \mathbb{R}^3$ es una base de vectores para $\mathbb{R}^3$ y $\eta \in \mathbb{R}^3$ es tal que $\eta = \lambda_1\xi_1 + \lambda_2\xi_2 + \lambda_3\xi_3$, entonces los coeficientes $\lambda_1, \lambda_2, \lambda_3$ están dados por las igualdades*

$$\lambda_1 = \frac{(\eta, \xi_2, \xi_3)}{(\xi_1, \xi_2, \xi_3)}, \; \lambda_2 = \frac{(\xi_1, \eta, \xi_3)}{(\xi_1, \xi_2, \xi_3)}, \; \lambda_3 = \frac{(\xi_1, \xi_2, \eta)}{(\xi_1, \xi_2, \xi_3)}$$

EJEMPLO. Sea $\eta = (-1, 3, -1) \in \mathbb{R}^3$ y sea la base de $\mathbb{R}^3$ dada por

$$\{\xi_1 = (1, 1, 0), \xi_2 = (0, 1, 1), \xi_3 = (1, 1, 1)\}$$

◁ Al calcular los determinantes, tenemos,

$$(\xi_1, \xi_2, \xi_3) = \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 1$$

$$(\eta, \xi_2, \xi_3) = \begin{vmatrix} \eta \\ \xi_2 \\ \xi_3 \end{vmatrix} = \begin{vmatrix} -1 & 3 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = -\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} + \begin{vmatrix} 3 & -1 \\ 1 & 1 \end{vmatrix} = 4$$

$$(\xi_1, \eta, \xi_3) = \begin{vmatrix} \xi_1 \\ \eta \\ \xi_3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ -1 & 3 & -1 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 3 & -1 \\ 1 & 1 \end{vmatrix} - \begin{vmatrix} -1 & -1 \\ 1 & 1 \end{vmatrix} = 4$$

$$(\xi_1, \xi_2, \eta) = \begin{vmatrix} \xi_1 \\ \xi_2 \\ \eta \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 3 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 3 & -1 \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ -1 & -1 \end{vmatrix} = -5$$

de donde, por el Lema 2.7 se tiene que

$$\lambda_1 = \frac{(\eta_1, \xi_1, \xi_2)}{(\xi_1, \xi_2, \xi_3)} = \frac{4}{1} = 4$$

$$\lambda_2 = \frac{(\xi_1, \eta, \xi_3)}{(\xi_1, \xi_2, \xi_3)} = \frac{4}{1} = 4$$

$$\lambda_3 = \frac{(\xi_1, \xi_2, \eta)}{(\xi_1, \xi_2, \xi_3)} = \frac{-5}{1} = -5$$

hacen que

$$\lambda_1\xi_1 + \lambda_2\xi_2 + \lambda_3\xi_3 = 4(1,1,0) + 4(0,1,1) - 5(1,1,1)$$

$$= (4,4,0) + (0,4,4) + (-5,-5,-5) = (4,8,4) + (-5,-5,-5)$$

$$= (-1,3,-1) = \eta. \quad \triangleright$$

EJEMPLO. Sean los vectores básicos $\{\xi_1 = (1,2),\ \xi_2 = (-2,1)\}$ de $\mathbb{R}^2$. Entonces dado $\eta = (-1,3)$ se tiene que

$$\eta = \lambda_1\zeta_1 + \lambda_2\xi_2$$

para los valores de λ_1 y λ_2 dados por

$$\lambda_1 = \frac{\begin{vmatrix} \eta \\ \xi_2 \end{vmatrix}}{\begin{vmatrix} \xi_1 \\ \xi_2 \end{vmatrix}} = \frac{\begin{vmatrix} -1 & 3 \\ -2 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ -2 & 1 \end{vmatrix}} = \frac{5}{5} = 1$$



$$\lambda_2 = \frac{\begin{vmatrix} \xi_1 \\ \eta \end{vmatrix}}{\begin{vmatrix} \xi_1 \\ \xi_2 \end{vmatrix}} = \frac{\begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ -2 & 1 \end{vmatrix}} = \frac{5}{5} = 1$$

Esto se comprueba directamente de la igualdad

$$\lambda_1\xi_1 + \lambda_2\xi_2 = 1(1,2) + 1(-2,-1) = (1,2) + (-2,1) = (-1,3) \quad \triangleright$$

Consideremos ahora una base $\{\xi_1, \xi_2, \xi_3\}$ de $\mathbb{R}^3$ y la ordenamos según el signo del determinante dado por el triple producto escalar

$$(\xi_1, \xi_2, \xi_3) = \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{vmatrix}$$

DEFINICIÓN. Decimos que la base dada $\{\xi_1, \xi_2, \xi_3\}$ de $\mathbb{R}^3$ tiene una **orientación positiva a la derecha** (**regla de la mano derecha**) si al considerar en el orden ξ_1, ξ_2, ξ_3, el triple producto escalar

$$(\xi_1, \xi_2, \xi_3)$$

es positivo.

EJEMPLO. Los vectores básicos

$$\{e_1 = (1,0,0),\ e_2 = (0,0,0),\ e_3 = (0,0,1)\}$$

dados en ese orden le dan una orientación a la derecha al espacio $\mathbb{R}^3$, en virtud de que

$$(e_1, e_2, e_3) = \begin{vmatrix} e_1 \\ e_2 \\ e_3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{vmatrix} = 1$$

EJEMPLO. En cambio, si se toma el orden de los vectores básicos canónicos dado por

$$\{e_1 = (1,0,0),\ e_3 = (0,0,1),\ e_2 = (0,1,0)\}$$

entonces tal orden le da una **orientación negativa** a la derecha a $\mathbb{R}^3$ en virtud de que

$$(e_1, e_3, e_2) = \begin{vmatrix} e_1 \\ e_2 \\ e_3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} = -1$$



EJEMPLO. El conjunto de vectores básicos

$$\{\xi_1 = (1,1,0),\ \xi_2 = (0,1,1),\ \xi_3 = (1,1,1)\}$$

ordenados así, le dan una orientación a $\mathbb{R}^3$ positiva a la derecha en virtud de la igualdad

$$(\xi_1, \xi_3, \xi_2) = \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 1$$

EJEMPLO. Los vectores básicos $\{\xi_1 = (1,2),\ \xi = (-2,1)\}$ de $\mathbb{R}^2$ le dan una orientación positiva a la derecha al plano en virtud de que el siguiente determinante es positivo

$$\begin{vmatrix} \xi_1 \\ \xi_2 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ -2 & 1 \end{vmatrix} = 5$$

Se observa que al tomar dos vectores no nulos arbitrarios ξ, η, sí $[\xi, \eta] \neq 0$, entonces los vectores unitarios

$$\left\{ \frac{\xi}{||\xi||}, \frac{\eta}{||\eta||}, \frac{[\xi,\eta]}{||[\xi,\eta]||} \right\}$$

conforman una **base** que está orientada positiva a la derecha (**regla de la mano derecha**) por la elección del producto $[\xi, \eta]$.

Todos los conceptos mencionados forman parte de los conocimientos básicos de los cursos de Geometría analítica del plano y del espacio.

2.5 Vectores y valores propios de una matriz

En algunos problemas de aparición cotidiana en las ciencias naturales y la ingeniería es necesario resolver una igualdad matricial del tipo

$$A\xi = \lambda\xi$$

donde A es una matriz cuadrada $(n \times n)$, ξ es un vector columna $(n \times 1)$ y λ es un escalar.

En otras palabras, dada un matriz cuadrada A, el problema es encontrar un vector ξ y un escalar λ que hagan válida la mencionada ecuación.

Discutiremos el caso de dimensión 3 y mencionaremos los resultados que se pueden extender para dimensiones arbitrarias.



Sea A una matriz de 3×3,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

y definimos el vector incógnita ξ por el vector columna,

$$\xi = \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix}$$

Entonces la ecuación matricial se escribe

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} = \lambda \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix}$$

o bien,

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} - \lambda \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix}$$

$$= \left[\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right] \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix}$$

que matricialmente se escribe

$$0 = (A - \lambda I)\xi$$

y que induce el sistema de ecuaciones lineales

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix}$$

es decir, el sistema homogéneo en ξ^1, ξ^2, ξ^3,

$$\begin{cases} (a_{11} - \lambda)\xi^1 + a_{12}\xi^2 + a_{13}\xi^3 = 0 \\ a_{21}\xi^1 + (a_{22} - \lambda)\xi^2 + a_{23}\xi^3 = 0 \\ a_{31}\xi^1 + a_{32}\xi^2 + (a_{33} - \lambda)\xi^3 = 0 \end{cases}$$

Tal sistema lineal homogéneo tiene asociada a la matriz $(A - \lambda I)$, y por el teorema 1.7, el sistema lineal homogéneo obtenido tiene soluciones no triviales ($\xi \neq (0,0,0)$), si y sólo sí, $\det(A - \lambda I) = 0$, donde det define la operación determinante de una matriz.

Caracterizamos a los escalares λ y a los vectores ξ que cumplen la condición matricial dada inicialmente, para la matriz A.

DEFINICIÓN. Si A es una matriz cuadrada real, entonces el escalar λ se dirá un **valor propio** de A si satisface la ecuación

$$\det(\lambda I - A) = 0$$

llamada la **ecuación característica** de la matriz A.

Observamos que, de la relación $0 = \det(\lambda I - A) = \det(A - \lambda I)$ para calcular los valores propios de una matriz, ocuparemos cualquiera de estas ecuaciones.

EJEMPLO. Calcule los valores propios de la matriz real de 2×2

$$A = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$$

$\lhd$ Al realizar las operaciones matriciales

$$\lambda I - A = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} \lambda - 3 & -2 \\ 1 & \lambda \end{pmatrix}$$

se obtiene,

$$\det(\lambda I - A) = \det \begin{pmatrix} \lambda - 3 & -2 \\ 1 & \lambda \end{pmatrix} = \lambda^2 - 3\lambda + 2$$

lo que implica que la ecuación característica de A es

$$\lambda^2 - 3\lambda + 2 = 0$$

Las raíces de esta ecuación son $\lambda = 1$ y $\lambda = 2$, y son los valores propios de A. $\rhd$

EJEMPLO. Calcule los valores propios de la matriz

$$A = \begin{pmatrix} -2 & -1 \\ 5 & 2 \end{pmatrix}$$

$\lhd$ Procediendo como en el ejemplo anterior se obtiene

$$\det(\lambda I - A) = \det \begin{pmatrix} \lambda + 2 & 1 \\ -5 & \lambda - 2 \end{pmatrix} = \lambda^2 + 1$$

Con lo cual, los valores propios de A deben satisfacer la ecuación cuadrática $\lambda^2 + 1 = 0$.

Dado que las únicas soluciones de esta ecuación son los **números imaginarios** $\lambda = i$ y $\lambda = -i$, se sigue A no tiene valores propios reales. ▷

EJEMPLO. Encuentre los valores propios de

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 3 & 2 & 4 \\ 0 & 0 & 4 \end{pmatrix}$$

Como en los ejemplos anteriores, se obtiene

$$\det(\lambda I - A) = \det \begin{pmatrix} \lambda - 2 & -1 & 0 \\ -3 & \lambda - 2 & 0 \\ 0 & 0 & \lambda - 4 \end{pmatrix} = \lambda^3 - 8\lambda^2 + 17\lambda - 4$$

Por lo tanto, los valores propios de A deben satisfacer la ecuación de tercer grado

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0$$

Por inspección $\lambda = 4$ es una raíz, es decir,

$$(\lambda - 4)(\lambda^2 - 4\lambda + 1) = 0$$

consiguientemente, las raíces restantes de tal cúbica satisfacen la ecuación

$$\lambda^2 - 4\lambda + 1 = 0$$

la cual se puede resolver mediante la fórmula cuadrática.

Por tanto, los valores propios de A son

$$\lambda = 4 \quad \lambda = 2 - \sqrt{3} \quad \text{y} \quad \lambda = 2 - \sqrt{3} \quad \triangleright$$

DEFINICIÓN. Sean, A una matriz cuadrada, y λ un valor propio de A. El vector ξ se dirá un **vector propio** de A asociado a λ si satisface la ecuación matricial

$$(A - \lambda I)\xi = 0$$

EJEMPLO. Sea la matriz cuadrada de 3×3 dada por

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

$\triangleleft$ Calculamos sus valores propios resolviendo la ecuación

$$0 = \det(A - \lambda I)$$

es decir,

$$0 = \begin{vmatrix} 2-\lambda & -1 & 0 \\ -1 & 2-\lambda & 0 \\ 0 & 0 & 4-\lambda \end{vmatrix} = (4-\lambda)(\lambda^2 - 4\lambda + 3)$$

cuyas raíces son $\lambda = 1, 3, 4$.

Para calcular los vectores propios asociados a cada valor propio, consideramos la ecuación matricial

$$(A - \lambda I)\xi = 0$$

para cada valor de λ, es decir, el sistema homogéneo,

$$\left\{ \begin{array}{c} (2-\lambda)\xi^1 - \xi^2 = 0 \\ -\xi^1 + (2-\lambda)\xi^2 = 0 \\ (4-\lambda)\xi^3 = 0, \end{array} \right. \qquad \xi = (\xi^1, \xi^2, \xi^3).$$

Para $\lambda = 1$ se tiene el sistema

$$\left\{ \begin{array}{c} \xi^1 - \xi^2 = 0 \\ -\xi^1 + \xi^2 = 0 \\ 3\xi^3 = 0 \end{array} \right.$$

Notamos que la primer ecuación es la misma que la segunda, y que necesariamente $\xi^3 = 0$. Si ponemos $x = t$, entonces la solución general de este sistema es

$$(\xi^1, \xi^2, \xi^3) = (t, t, 0) = t(1, 1, 0)$$

lo que indica que un vector propio asociado a $\lambda = 1$ es

$$\xi_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

Para $\lambda = 3$ se obtiene el sistema homogéneo

$$\left\{ \begin{array}{c} -\xi^1 - \xi^2 = 0 \\ -\xi^1 - \xi^2 = 0 \\ \xi^3 = 0 \end{array} \right.$$

cuya solución general es del tipo

$$(\xi^1, \xi^2, \xi^3) = (t, -t, 0) = t(1, -1, 0)$$



lo cual implica que un vector propio asociado a $\lambda = 3$ es

$$\xi_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Análogamente, para $\lambda = 4$ se tiene el sistema lineal

$$\begin{cases} -2\xi^1 - \xi^2 = 0 \\ -\xi^1 - 2\xi^2 = 0 \\ 0\xi^3 = 0 \end{cases}$$

La tercer ecuación nos dice que ξ^3 es arbitraria, mientras que el sistema de 2×2 formado por las dos primeras ecuaciones tiene sólo la solución trivial para ξ^1, ξ^2, debido a que

$$\begin{vmatrix} -2 & -1 \\ -1 & -2 \end{vmatrix} = 3 \neq 0$$

Por lo tanto, la solución general de este último sistema homogéneo es

$$(\xi^1, \xi^2, \xi^3) = (0, 0, t) = t(0, 0, 1)$$

lo que implica que un vector propio asociado a $\lambda = 4$ es

$$\xi_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

De esta manera, el conjunto

$$\{\xi_1 = (1, 1, 0),\ \xi_2 = (1, 1, 0),\ \xi_3 = (0, 0, 1)\}$$

es un conjunto de vectores propios de A.

Observamos que este conjunto es una base de $\mathbb{R}^3$ debido a que

$$(\xi_1, \xi_2, \xi_3) = \begin{vmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = -2 \neq 0 \quad \triangleright$$

EJEMPLO. Sea la matriz

$$A = \begin{pmatrix} 1 & 2 & -4 \\ 2 & -2 & -2 \\ -4 & -2 & 1 \end{pmatrix}$$

Calculamos sus vectores y valores propios.

◁ Calculamos primeramente el determinante

$$0 = \begin{vmatrix} 1-\lambda & 2 & -4 \\ 2 & -2-\lambda & -2 \\ -4 & -2 & 1-\lambda \end{vmatrix} = \lambda^3 - 27\lambda - 54$$

La ecuación cúbica obtenida tiene raíces enteras $\lambda_1 = 6$, $\lambda_2 = -3, \lambda_3 = -3$, donde una de ellas es doble.

El sistema lineal homogéneo para un valor propio λ es

$$\begin{cases} (1-\lambda)\xi^1 + 2\xi^2 - 4\xi^3 = 0 \\ 2\xi^1 + (-2-\lambda)\xi^2 - 2\xi^3 = 0 \\ -4\xi^1 - 2\xi^2 + (1-\lambda)\xi^3 = 0, \end{cases} \qquad \xi = (\xi^1, \xi^2, \xi^3).$$

Para el valor propio doble $\lambda = -3$ se tiene el sistema lineal homogéneo en las incógnitas ξ^1, ξ^2, ξ^3.

$$\begin{cases} 4\xi^1 + 2\xi^2 - 4\xi^3 = 0 \\ 2\xi^1 + \xi^2 - 2\xi^3 = 3 \\ -4\xi^1 - 2\xi^2 + 4\xi^3 = 0 \end{cases}$$

Al resolver por El método de Gauss-Jordan, se tiene,

$$\begin{pmatrix} 4 & 2 & -4 & 0 \\ 2 & 1 & -2 & 0 \\ -4 & -2 & 4 & 0 \end{pmatrix} (R_1 - 2R_2) \to R_2, (R_1 - R_3) \to R_3$$

$$\sim \begin{pmatrix} 4 & 2 & -4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

que induce la única ecuación con tres incógnitas

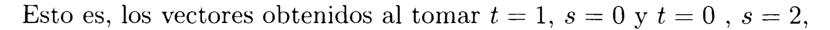
$$4\xi^1 + 2\xi^2 - 4\xi^3 = 0 \iff 2\xi^1 + \xi^2 - 2\xi^3 = 0$$

$$\iff \xi^3 = \xi^1 + \frac{1}{2}\xi^2$$

Al poner $\xi^1 = t$, $\xi^2 = s$, la solución general tiene la forma,

$$(\xi^1, \xi^2, \xi^3) = \left(t, s, t + \frac{1}{2}s\right) = (t, 0, t) + \left(0, s, \frac{1}{2}s\right)$$

$$= t(1, 0, 1) + s\left(0, 1, \frac{1}{2}\right)$$



Esto es, los vectores obtenidos al tomar $t = 1$, $s = 0$ y $t = 0$, $s = 2$,

$$\xi_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \xi_2 = 2\begin{pmatrix} 0 \\ 1 \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

son vectores propios asociados a $\lambda = -3$.

Para el valor propio $\lambda = 6$ se tiene el sistema homogéneo

$$\begin{cases} -5\xi^1 + 2\xi^2 - 4\xi^3 = 0 \\ 2\xi^1 - 8\xi^2 - 2\xi^3 = 0 \\ -4\xi^1 - 2\xi^2 - 5\xi^3 = 0 \end{cases}$$

que se resuelve mediante El método de Gauss-Jordan de la siguiente forma.

$$\begin{pmatrix} -5 & 2 & -4 & 0 \\ 2 & -8 & -2 & 0 \\ -4 & -2 & -5 & 0 \end{pmatrix} (2R_1 + 5R_2) \to R_2,\ (4R_1 - 5R_3) \to R_3$$

$$\sim \begin{pmatrix} -5 & 2 & -4 & 0 \\ 0 & -36 & -18 & 0 \\ 0 & 18 & 9 & 0 \end{pmatrix} (R_2 + 2R_3) \to R_3$$

$$\sim \begin{pmatrix} -5 & 2 & -4 & 0 \\ 0 & -36 & -18 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \left(\frac{R_2}{18}\right) \to R_2$$

$$\sim \begin{pmatrix} -5 & 2 & -4 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Es decir, se tiene el sistema,

$$\begin{cases} -5\xi^1 + 2\xi^2 - 4\xi^3 = 0 \\ 2\xi^2 - \xi^3 = 0 \end{cases}$$

en el cual, si se pone $\xi^2 = t$ se tiene una solución general,

$$(\xi^1, \xi^2, \xi^3) = \left(\frac{6}{5}t, t, 2t\right) = t\left(\frac{6}{5}, 1, 2\right)$$

De esta forma, tomando $t = 5$, se tiene el vector

$$\xi_3 = 5\begin{pmatrix} 6/5 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 5 \\ 10 \end{pmatrix}$$

que es un vector propio asociado a $\lambda = 6$.

Notamos que el sistema de vectores propios

$$\{\xi_1 = (1,0,1),\ \xi_2 = (0,2,1),\ \xi_3 = (5,6,10)\}$$

es una base de vectores para $\mathbb{R}^3$, debido a que

$$(\xi_1, \xi_2, \xi_3) = \begin{vmatrix} 1 & 0 & 5 \\ 0 & 2 & 6 \\ 1 & 1 & 10 \end{vmatrix} = -24 \neq 0 \quad \triangleright$$

EJEMPLO. Consideremos la matriz

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

y procedamos a calcular sus vectores y valores propios.

$\triangleleft$ Un cálculo directo nos muestra que

$$0 = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 1 & 2-\lambda & 0 \\ 0 & 1 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda)^2$$

lo que nos indica inmediatamente que los valores propios son $\lambda_1 = 1$, $\lambda_2 = 2$ y $\lambda_3 = 2$.

El sistema de ecuaciones lineales homogéneas que nos ayuda a calcular los vectores propios es

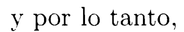
$$\begin{cases} (1-\lambda)\xi^1 = 0 \\ \xi^1 + (2-\lambda)\xi^2 = 0 \\ \xi^2 + (2-\lambda)\xi^3 = 0 \end{cases}$$

De esta forma, para el valor propio $\lambda = 1$ se tiene el sistema lineal,

$$\begin{cases} 0\xi^1 = 0 \\ \xi^1 + \xi^2 = 0 \\ \xi^2 + \xi^3 = 0 \end{cases}$$

lo que nos indica que ξ^1 es arbitrario y que $\xi^2 = -\xi^1$, $\xi^3 = \xi^1$. En otras palabras, la solución general es

$$(\xi^1, \xi^2, \xi^3) = (t, -t, t) = t(1, -1, 1)$$



y por lo tanto,

$$\xi_1 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

es un vector propio asociado a $\lambda = 1$.

Para el caso $\lambda = 2$ se tiene el sistema lineal

$$\begin{cases} -\xi^1 = 0 \\ \xi^1 + 0\xi^2 = 0 \\ \xi^2 + 0\xi^3 = 0 \end{cases} \iff \begin{cases} \xi^1 = 0 \\ 0\xi^2 = 0 \\ \xi^2 + 0\xi^3 = 0 \end{cases}$$

lo que nos indica, de la primera ecuación que $\xi^1 = 0$, de la tercera que $\xi^2 = 0$ y que ξ^3 es arbitrario, es decir, la solución del sistema es del tipo general

$$(\xi^1, \xi^2, \xi^3) = (0, 0, t) = t(0, 0, 1)$$

De esta manera, el vector

$$\xi_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

es el único vector propio asociado a $\lambda = 2$. ▷

De la discusión inicial se tiene el siguiente resultado.

TEOREMA 2.2 *Las siguientes proposiciones son equivalentes para una matriz cuadrada $(n \times n)$ dada.*

a. *A tiene un valor propio λ.*

b. $\det(A - \lambda I) = 0$.

c. *Existe un vector propio ξ para A asociado al valor propio λ.*

Una matriz cuadrada se dice **diagonal** si sus entradas que no están sobre su diagonal son todas idénticamente nulas.

EJEMPLO. La matriz cuadrada

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

es una matriz diagonal.

DEFINICIÓN. Una matriz cuadrada real $(n \times n)$ A se dice **diagonalizable** si existe una matriz invertible B tal que el producto $B^{-1}AB$ es una matriz diagonal. En este caso se dice que B diagonaliza a la matriz A.

Tenemos el siguiente resultado para el proceso de diagonalización de una matriz cuadrada.

TEOREMA 2.3 *Para una matriz cuadrada A de $(n \times n)$ son equivalentes las siguientes proposiciones.*

a. *A es diagonalizable*

b. *Los vectores propios A conforman una base de $\mathbb{R}^n$ y la matriz B formada por los vectores propios de A (como vectores columna) diagonaliza A.*

$\lhd$ Daremos la prueba apenas para las matrices de 3×3, indicando que para dimensiones mayores la prueba es análoga.

Suponemos que $A = A_{3\times 3}$ es diagonalizable. Entonces existe una matriz

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

tal que $B^{-1}\ AB$ es una matriz diagonal $D = B^{-1}AB$ que tiene forma,

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

Esto implica necesariamente que $BD = AB$.

Por un lado, se tiene

$$BD = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = \begin{pmatrix} \lambda_1 b_{11} & \lambda_2 b_{12} & \lambda_3 b_{13} \\ \lambda_1 b_{21} & \lambda_2 b_{22} & \lambda_3 b_{23} \\ \lambda_1 b_{31} & \lambda_2 b_{32} & \lambda_3 b_{33} \end{pmatrix}$$

$$= (\lambda_1 b_1 \quad \lambda_2 b_2 \quad \lambda_3 b_3)$$

donde hemos escrito a la última matriz mediante sus columnas.

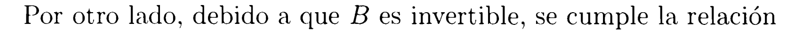
Por otro lado, si escribimos a la matriz B mediante sus columnas, de la multiplicación de matrices se tiene

$$AB = A(b_1 \quad b_2 \quad b_3) = (Ab_1 \quad Ab_2 \quad Ab_2)$$

De esta manera, al igualar columnas en la igualdad $AB = BD$ se obtienen las igualdades

$$Ab_1 = \lambda_1 b_1, \quad Ab_2 = \lambda_2 B_2, \quad Ab_3 = \lambda_3 b_3$$

es decir, $\{b_1, b_2, b_3\}$ es un conjunto de vectores propios de A asociados a los valores propios correspondientes $\lambda_1, \lambda_2, \lambda_3$.



Por otro lado, debido a que B es invertible, se cumple la relación

$$(b_1, b_2, b_3) = \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} \neq 0$$

que hace de $\{b_1, b_2, b_3\}$ una base de vectores propios para $\mathbb{R}^3$. Esto demuestra que **a.** implica **b.**

Para probar la recíproca, sean $\{\xi_1, \xi_2, \xi_3\}$ los vectores propios de A asociados a los valores propios λ_1, λ_2 y λ_3 respectivamente, y sea

$$B = (\xi_1 \quad \xi_2 \quad \xi_3) = \begin{pmatrix} \xi_1^1 & \xi_1^2 & \xi_1^3 \\ \xi_2^1 & \xi_2^2 & \xi_2^3 \\ \xi_3^1 & \xi_3^2 & \xi_3^3 \end{pmatrix}$$

Entonces B es invertible pues los vectores son linealmente independientes.

Por las consideraciones anteriores se tiene que

$$AB = A(\xi_1 \quad \xi_2 \quad \xi_3) = (A\xi_1 \quad A\xi_2 \quad A\xi_3)$$

$$= (\lambda_1\xi_1 \quad \lambda_2\xi_2 \quad \lambda_3\xi_3) = (\xi_1 \quad \xi_2 \quad \xi_3)\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = BD$$

donde D es la matriz diagonal, cuyas entradas diagonales son los valores propios de A, es decir,

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

De esta manera, B diagonaliza a A pues se cumple la igualdad matricial $AB = BD$, o equivalentemente, $B^{-1}AB = D$. Esto termina la prueba del teorema. $\triangleright$

Otro resultado útil para el proceso se diagonalización es el siguiente, cuya prueba omitimos por ser más técnica.

LEMA 2.8 *Si los valores propios de una matriz cuadrada A son todos diferentes entre sí, entonces los vectores propios de A conforman una base de vectores propios.*

Este resultado, junto con el teorema 2.3, implica el siguiente que es de gran utilidad.

TEOREMA 2.4 *Si la matriz cuadrada A de $n \times n$ tiene diferentes entre sí a todos sus valores propios, entonces es diagonalizable.*

Esto implica necesariamente que $BD = AB$. $\triangleleft$ Por el Lema 2.8, si los valores propios de A son diferentes entre sí, entonces los vectores propios de A conforman una base de $\mathbb{R}^n$. Por el teorema 2.3, A es necesariamente diagonalizable $\triangleright$

EJEMPLO. Para la matriz cuadrada

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

se obtuvieron los vectores propios

$$\{\xi_1 = (1,1,0), \xi_2 = (1,-1,0), \xi_2 = (0,0,1)\}$$

asociados a los valores propios $\lambda = 1, \lambda_3 = 3$ y $\lambda = 4$ respectivamente, y que conforman una base de $\mathbb{R}^3$.

Sea B la matriz

$$B = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

que cambia la base canónica $\{e_1, e_2, e_3\}$ por la base de vectores propios $\{\xi_1, \xi_2,$
$\xi_3\}$.

Por el teorema 1.5 se tiene que

$$B^{-1} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

y un cálculo directo prueba que

$$B^{-1}AB = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1/2 & 1/2 & 0 \\ 3/2 & -3/2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

es una matriz diagonal, con los valores propios como elementos diagonales.

Consecuentemente A es diagonalizable por la matriz B. $\triangleright$



El siguiente ejemplo nos muestra que una matriz A puede ser diagonalizable, aún sin tener todos sus valores propios distintos.

EJEMPLO. Para la matriz cuadrada

$$A = \begin{pmatrix} 1 & 2 & -4 \\ 2 & -2 & -2 \\ -4 & -2 & 1 \end{pmatrix}$$

se obtuvo una base de vectores propios

$$\{\xi_1 = (1, 0, 1), \xi_2 = (0, 2, 1), \xi_3 = (5, 6, 10), \}$$

asociados a los vectores propios $\lambda_1 = \lambda_2 = -3$ y $\lambda_3 = 6$ respectivamente.

Sea la matriz B dada por

$$B = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 2 & 6 \\ 1 & 1 & 10 \end{pmatrix}$$

que cambia la base canónica $\{e_1, e_2, e_3\}$ por la base de vectores propios obtenida.

Mediante El método de Gauss-Jordan de la sección 1.8 se puede obtener que

$$B^{-1} = \begin{pmatrix} 7/2 & 5/4 & -5/2 \\ 3/2 & 5/4 & -3/2 \\ -1/2 & -1/4 & 1/2 \end{pmatrix}$$

y un cálculo directo demuestra que

$$B^{-1}AB = \begin{pmatrix} 7/2 & 5/4 & -5/2 \\ 3/2 & 5/4 & -3/2 \\ -1/2 & -1/4 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & 2 & -4 \\ 2 & -2 & -2 \\ -4 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 5 \\ 0 & 2 & 6 \\ 1 & 1 & 10 \end{pmatrix}$$

$$= \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

En otras palabras, la matriz B diagonaliza a la matriz A. $\triangleright$

EJEMPLO. Al considerar la matriz

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

se obtuvieron los valores propios $\lambda_1 = 1$ y $\lambda_2 = 2$, y apenas dos vectores propios correspondientes a tales valores propios

$$\{\xi_1 = (1, -1, -1),\ \xi_2 = (0, 0, 1)\}$$

$\lhd$ Como este conjunto es linealmente independiente, pero no conforma una base de $\mathbb{R}^3$, se sigue que la matriz A no es diagonalizable. $\rhd$

Sobre métodos más generales de diagonalización el lector interesado puede referirse al trabajo de **Anton** (1981), donde encontrará más datos.

2.6 Rectas y Planos en $\mathbb{R}^3$

En esta parte introducimos los conceptos de recta y plano como objetos contenidos en un espacio tridimensional, caracterizándolos mediante los elementos que les definen.

Rectas en el espacio

Dados, un punto $p \in \mathbb{R}^n$ y un vector ξ, los puntos q de la recta $L \subset \mathbb{R}^n$ que pasa por p y es paralela a ξ satisfacen la ecuación

$$q = p + t\xi$$

para algún número real $t \in \mathbb{R}$ (véase la figura 2.11).

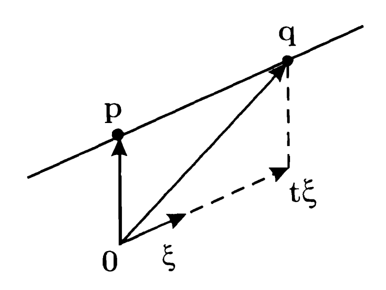


Figura 2.11: La recta en $\mathbb{R}^n$.

DEFINICIÓN. Decimos que la ecuación

$$q = p + t\xi, \quad t \in \mathbb{R}$$

es la **ecuación de la recta** contenida $\mathbb{R}^n$ por el punto p con el **vector director** ξ. Se llama también **ecuación paramétrica** de L pues depende del parámetro t.



EJEMPLO. Hallar la ecuación paramétrica de la recta L en el plano que pasa por los puntos $p_1 = (3, 2)$ y $p_2 = (4, 1)$.

$\lhd$ La recta pasa por $p = (4, 1)$ y la dirige el vector de posición

$$\xi = p_2 - p_1 = (4, 1) - (3, 2) = (1, -1)$$

De esta manera, si $q = (x, y)$ está sobre la recta, entonces satisface la relación

$$\begin{aligned}(x, y) &= (4, 1) + t(1, -1) = (4, 1) + (t, -t)\\ &= (4 + t, 1 - t)\end{aligned}$$

En otras palabras, el par de ecuaciones

$$\left\{\begin{array}{l} x = 4 + t \\ y = 1 - t \end{array}\right. , \quad t \in \mathbb{R}$$

definen a la recta L.

Si en las relaciones anteriores despejamos t, $t = x-4$, $t = 1-y$, entonces, al igualar tales ecuaciones se obtiene

$$x - 4 = 1 - y$$

que implica la igualdad

$$y = -x + 5$$

que también define a la recta L, pero sin parámetro (véase la figura 2.12) $\rhd$

EJEMPLO. Hallar la ecuación de la recta L en $\mathbb{R}^4$ que pasa por los puntos

$$p_1 = (2, 3, 4, 2), \quad p_2 = (6, 4, -5, 1)$$

$\lhd$ Tal recta pasa por $p = (2, 4, 3, 2)$ y tiene vector director

$$\xi = p_1 - p_2 = (2, 4, 3, 2) - (6, 4, -5, 1) = (-4, 0, 8, 1)$$

Por lo tanto, en las coordenadas (x, y, z, w) se tiene que todo punto $q = (x, y, z, w) \in L$, cumple las igualdades

$$\begin{aligned}(x, y, z, w) &= (2, 4, 3, 2) + t(-4, 0, 8, 1) = (2, 4, 3, 2) + (-4t, 0, 8t, t)\\ &= (2 - 4t, 4, 3 + 8t, 2 + t)\end{aligned}$$

Esto es, la recta L se define por el sistema paramétrico

$$\left\{\begin{array}{l} x = 2 - 4t \\ y = 4 \\ z = 3 + 8t \\ w = 2 + t \end{array}\right. \quad t \in \mathbb{R}$$

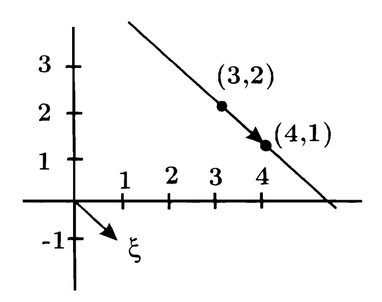


Figura 2.12: Recta $y = -x + 5$

Para ilustrar, evaluamos en estas igualdades algunos argumentos temporales $t \in \mathbb{R}$.

Si se toma $t = 0$ entonces el punto $(x = 2, y = 4, z = 3, w = 2)$ está en la recta L.

Si $t = -1$, entonces también el punto con coordenadas $x = 6, y = 4, z = -5$, $w = 1$ pertenece a L. $\triangleright$

En particular, aquí estamos interesados en las rectas contenidas en $\mathbb{R}^3$, tal que si pasan por un punto $p = (a, b, c)$ con vector director $\xi = (\xi^1, \xi^2, \xi^3)$, se tenga su definición mediante un sistema que tenga la forma,

$$\begin{cases} x = \xi^1 t + a \\ y = \xi^2 t + b \\ z = \xi^3 t + c \end{cases}$$

Planos en el espacio

Dados, un punto $p \in \mathbb{R}^3$ y un vector $n \in \mathbb{R}^3$, se define la **ecuación del plano** P pasando por el punto p con el **vector normal** (ortogonal) n, como el conjunto de puntos q (variables) $\in \mathbb{R}^3$, tales que satisfacen la ecuación.

$$< n, q - p >= 0$$

o bien, la ecuación equivalente,

$$< n, q >=< n, p >$$

La figura 2.13 ilustra la definición para el plano 2-dimensional en $\mathbb{R}^3$.

Notamos que para el caso de $\mathbb{R}^n$, se pueden hacer definiciones análogas para un $(n-1)$-plano.



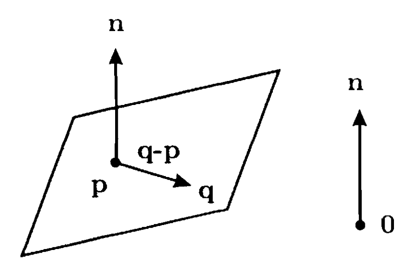


Figura 2.13: El plano en $\mathbb{R}^3$.

EJEMPLO. Sea el punto $p = (-3, 1, 2)$ en $\mathbb{R}^3$ y sea $n = (-4, 5, 3)$ un vector en $\mathbb{R}^3$. La ecuación del plano por el punto p y normal n es, en coordenadas $q = (x, y, z)$, viene dado por la igualdad

$$\lhd \quad < (-4, 5, 3), (x, y, z) >=< (-4, 5, 3), (-3, 1, -2) >$$

o equivalentemente,

$$-4x + 5y + 3z = 12 + 5 - 6 = 11$$

lo cual implica que la ecuación buscada del plano en $\mathbb{R}^3$ es

$$-4x + 5y + 3z = 11 \quad \rhd$$

En general, para la ecuación

$$< n, q >=< n, p >$$

del plano P pasando por $p = (p_1, p_2, p_3)$ con vector normal $n = (a, b, c)$ se tiene que en las coordenadas (x, y, z) los puntos $q = (x, y, z)$ de P satisfacen la ecuación

$$< (a, b, c), (x, y, z) >=< (a, b, c) , (p_1, p_2, p_3) >$$

o equivalentemente, satisfacen la ecuación

$$ax + by + cz = ap_1 + bp_2 + cp_3 = d$$

De esta forma, una ecuación del tipo

$$ax + by + cz = d$$

representa la ecuación de un plano en $\mathbb{R}^3$ conteniendo un vector normal $n = (a, b, c)$.



EJEMPLO. La ecuación

$$2x + 3y + 6z = 1$$

representa a un plano en $\mathbb{R}^3$ con un vector normal $n = (2, 3, 6)$

◁ Si hacemos $y = 0$, $z = 0$ sobre el plano $\mathbb{R}^2_{y,z}$ entonces, de la ecuación dada se tiene $2x = 1$, lo que implica que $x = \frac{1}{2}$.

Por otro lado, si hacemos $y = 0$, $x = 0$ sobre el plano $\mathbb{R}^2_{x,y}$ entonces se obtendría que en este caso $z = \frac{1}{6}$

Finalmente, si tomamos $x = 0$, $z = 0$ sobre el plano $\mathbb{R}^2_{x,z}$, entonces, a partir de la ecuación del plano se tiene que $y = \frac{1}{3}$.

Lo anterior nos da los cortes del plano sobre los ejes, y la extensión del triángulo mostrado en la figura 2.14 nos realiza totalmente el plano. ▷

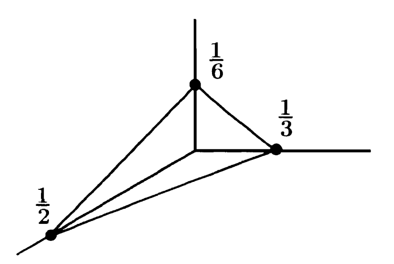


Figura 2.14: Parte del plano $2x + 3y + 6z = 1$ en el primer octante.

OBSERVACIÓN. Dos planos P_1, P_2 en el espacio son paralelos si sus respectivos vectores normales son paralelos.

EJEMPLO. Sean los planos en $\mathbb{R}^3$, dados por las ecuaciones

$$-x + 2y - 4z = 2, \quad 3x - 6y + 12z = 7$$

◁ Sus normales respectivos son

$$n_1 = (-1, 2, -4), \quad \text{y} \quad n_2 = (3, -6, 12)$$

Al calcular el producto vectorial de estos normales,

$$[n_1, n_2] = \begin{vmatrix} e_1 & e_2 & e_3 \\ -1 & 2 & -4 \\ 3 & -6 & 12 \end{vmatrix}$$

$$= e_1(24 - 24) - e_2(-12 + 12) + e_3(6 - 6) = (0, 0, 0)$$



se concluye que n_1 y n_2 son paralelos, lo que implica que los planos son paralelos. $\triangleright$

Generalizamos la idea de ángulo formado entre dos planos.

DEFINICIÓN. Sean dos planos P_1, P_2 contenidos en $\mathbb{R}^3$ y sean n_1, n_2 sus normales respectivos. Se define el **ángulo** θ formado por los planos P_1 y P_2 como el ángulo θ formado por sus normales n_1 y n_2. La figura 2.15 ilustra esta definición.

EJEMPLO. Dados los planos en $\mathbb{R}^3$

$$x + 2y + 3z = 1, \quad x - 2y + z = 3$$

verificar si son paralelos. De no serlo, calcular su intersección y el ángulo entre ellos.

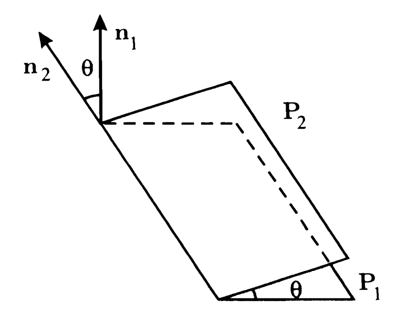


Figura 2.15: Ángulo formado por dos planos en $\mathbb{R}^3$.

$\triangleleft$ Los vectores normales respectivos son

$$n_1 = (1,2,3) \quad \text{y} \quad n_2 = (1,-2,1)$$

con lo que tenemos,

$$[n_1, n_2] = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 2 & 3 \\ 1 & -2 & 1 \end{vmatrix} = (2+6, -(1-3), -2-2)$$

$$= (8, 2, -4) \neq (0,0,0)$$

Por lo tanto, los normales n_1 y n_2 no son paralelos y consecuentemente los planos se intersectan. Los puntos (x, y, z) contenidos en la intersección deberán satisfacer ambas ecuaciones.



Tales puntos se determinan al resolver el sistema lineal 2×3,

$$\begin{cases} x + 2y + 3z = 1 \\ x - 2y + z = 3 \end{cases}$$

el cual, al hacer $z = t$ induce el sistema,

$$\begin{cases} x + 2y = 1 - 3t \\ x - 2y = 3 - t \end{cases}$$

Utilizamos La regla de Cramer para resolver este sistema. Calculando los determinantes, obtenemos,

$$D = \begin{vmatrix} 1 & 2 \\ 1 & -2 \end{vmatrix} = -4$$

$$D_x = \begin{vmatrix} 1-3t & 2 \\ 3-t & -2 \end{vmatrix} = -2(1-3t) - 2(3t) = 8t - 8$$

$$D_y = \begin{vmatrix} 1 & 1-3t \\ 1 & 3-t \end{vmatrix} = (3-t) - (1-3t) = 2t + 2$$

y de esta manera,

$$x = \frac{Dx}{D} = \frac{8t-8}{-4} = -2t + 2$$

$$y = \frac{Dy}{D} = \frac{2t+2}{-4} = -\frac{1}{2}t - \frac{1}{2}$$

Por lo tanto, la solución del sistema (intersección de los planos) es la recta parametrizada

$$\begin{cases} x = -2t + 2 \\ y = -\frac{1}{2}t - \frac{1}{2} \\ z = t \end{cases}$$

donde $t \in \mathbb{R}$ es un parámetro temporal.

En particular, si $t = 0$ se tiene la solución particular $x = 2, y = -\frac{1}{2}, z = 0$, es decir, el punto $\left(2, \frac{1}{2}, 0\right)$ está en ambos planos, lo que se verifica directamente al sustituir en las ecuaciones de los planos,

$$\begin{cases} 2 + 2\left(-\frac{1}{2}\right) + 3(0) = 1 \\ 2 - 2\left(-\frac{1}{2}\right) + 0 = 3 \end{cases}$$

Para calcular el ángulo entre los planos se calcula el ángulo θ entre sus vectores normales

$$\theta = \arccos\left(\frac{< n_1, n_2 >}{||n_1|| ||n_2||}\right) = \arccos\left(\frac{1 - 4 + 3}{\sqrt{14}\sqrt{6}}\right)$$



es decir,

$$\theta = \arccos(0) = \frac{\pi}{2}$$

lo que nos dice que los planos son ortogonales. $\triangleright$

Es fácil probar el siguiente Lema utilizando las ideas del ejemplo anterior.

LEMA 2.9 *Si la pareja de planos*

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases}$$

no es paralela, entonces se intersectan en una recta parametrizada. Más aún, forman el ángulo θ dado por,

$$\theta = \arccos\left(\frac{< n_1, n_2 >}{||n_1||\,||n_2||}\right)$$

donde $n_1 = (a_1, b_1, c_1)$, $n_2 = (a_2, b_2, c_2)$ son los normales respectivos.

$\triangleleft$ Es un cálculo directo al resolver el sistema de ecuaciones lineales dado por los planos. $\triangleright$

EJEMPLO. Calcular la intersección de los planos en $\mathbb{R}^3$,

$$\begin{cases} x - y + 3z = 9 \\ 3x - 5y + z = -4 \\ 4x - 7y + z = 5 \end{cases}$$

$\triangleleft$ Resolvemos el sistema de ecuaciones de 3×3 usando La regla de Cramer.

$$D = \begin{vmatrix} 1 & -1 & 3 \\ 3 & -5 & 1 \\ 4 & -7 & 1 \end{vmatrix} = \begin{vmatrix} -5 & 1 \\ -7 & 1 \end{vmatrix} + \begin{vmatrix} 3 & 1 \\ 4 & 1 \end{vmatrix} + 3\begin{vmatrix} 3 & -5 \\ 4 & -7 \end{vmatrix} = -2$$

$$D_x = \begin{vmatrix} 9 & -1 & 3 \\ -4 & -5 & 1 \\ 5 & -7 & 1 \end{vmatrix} = 3\begin{vmatrix} -4 & -5 \\ 5 & -7 \end{vmatrix} - \begin{vmatrix} 9 & -1 \\ 5 & -7 \end{vmatrix} + \begin{vmatrix} 9 & -1 \\ -4 & -5 \end{vmatrix} = 168$$

$$D_y = \begin{vmatrix} 1 & 9 & 3 \\ 3 & -4 & 1 \\ 4 & 5 & 1 \end{vmatrix} = 3\begin{vmatrix} 3 & -4 \\ 4 & 5 \end{vmatrix} - \begin{vmatrix} 1 & 9 \\ 4 & 5 \end{vmatrix} + \begin{vmatrix} 1 & 9 \\ 3 & -4 \end{vmatrix} = 93$$



$$D_z = \begin{vmatrix} 1 & -1 & 9 \\ 3 & -5 & -4 \\ 4 & -7 & 5 \end{vmatrix} = \begin{vmatrix} -5 & 1 \\ -7 & 5 \end{vmatrix} + \begin{vmatrix} 3 & -4 \\ 4 & 5 \end{vmatrix} + 9 \begin{vmatrix} 3 & -5 \\ 4 & -7 \end{vmatrix} = -7$$

Por lo tanto, las soluciones al sistema son

$$x = \frac{D_x}{D} = \frac{168}{-2} = -84, \quad y = \frac{D_y}{D} = \frac{93}{-2} = -\frac{93}{2}, \quad z = \frac{D_z}{D} = \frac{-7}{-2} = \frac{7}{2}$$

De esta manera, el punto donde se intersectan los planos tiene coordenadas

$$p = \left(-84, -\frac{93}{2}, \frac{7}{2}\right) \quad \triangleright$$

De igual forma, para dimensiones mayores ($n \geq 3$) podemos definir la ecuación de un plano en $\mathbb{R}^n$ (**hiperplano**) mediante los mismos elementos, n vector normal, p punto por donde pasa y q el vector variable, utilizando la ecuación

$$< n, p > = < n, q >$$

Por ejemplo, la ecuación

$$6x - 3y + z - 2w + 3v = 8$$

define un plano de dimensión 4 en el espacio $\mathbb{R}^5$ en el sistema de coordenadas (x, y, z, v, w) con un vector normal $n = (6, -3, 1, -3, -2)$.

EJEMPLO. Calcular la intersección en $\mathbb{R}^4$ de los **hiperplanos** de dimensión 3, definidos por,

$$\begin{cases} 2x + 7y + 3z + t = 5 \\ x + 3y + 5z - 2t = 3 \\ z + 5y - 9z + 8z = 1 \\ 5x + 18y + 4z + 5t = 12 \end{cases}$$

$\triangleleft$ Para calcular la intersección es necesario resolver el sistema, y lo realizamos mediante El método de Gauss-Jordan.

$$\begin{pmatrix} 2 & 7 & 3 & 1 & 5 \\ 1 & 3 & 5 & -2 & 3 \\ 1 & 5 & -9 & 8 & 1 \\ 5 & 18 & 4 & 5 & 12 \end{pmatrix} \quad (R_1 \leftrightarrow R_2)$$

$$\sim \begin{pmatrix} 1 & 3 & 5 & -2 & 3 \\ 2 & 7 & 3 & 1 & 5 \\ 1 & 5 & -9 & 8 & 1 \\ 5 & 18 & 4 & 5 & 12 \end{pmatrix} \quad (2R_1 - R_2) \rightarrow R_2,$$

$$(R_1 - R_3) \to R_3,\ (5R_1 - R_4) \to R_4$$

$$\sim \begin{pmatrix} 1 & 3 & 5 & -2 & 3 \\ 0 & -1 & 7 & -5 & 1 \\ 0 & -2 & 14 & -10 & 2 \\ 0 & -3 & 21 & -15 & 3 \end{pmatrix} \quad \left(\frac{R_3}{2}\right) \to R_3,\ \left(\frac{R_4}{-3}\right) \to R_4$$

$$\sim \begin{pmatrix} 1 & 3 & 5 & -2 & 3 \\ 0 & -1 & 7 & -5 & 1 \\ 0 & 1 & -7 & 5 & -1 \\ 0 & 1 & -7 & 5 & -1 \end{pmatrix} \quad (R_3 + R_2) \to R_3,\ (R_4 + R_2) \to R_4$$

$$\sim \begin{pmatrix} 1 & 3 & 5 & -2 & 3 \\ 0 & -1 & 7 & -5 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

De esta manera, nos queda el sistema lineal de 2×4 dado por

$$\begin{cases} x + 3y + 5z - 2t = 3 \\ -y + 7z - 5t = 1 \end{cases},$$

Tal sistema lo resolvemos por La regla de Cramer en virtud de que

$$D = \begin{vmatrix} 1 & 3 \\ 0 & -1 \end{vmatrix} = -1$$

lo que nos asegura que el sistema

$$\begin{cases} x + 3y = 3 - 5z + 2t \\ -y = 1 - 7z + 5t \end{cases}$$

tiene solución en dos variables libres z, t.

La segunda ecuación indica la solución para la variable y, y es suficiente con calcular sólo el siguiente determinante,

$$D_x = \begin{vmatrix} 3 - 5z + 2t & 3 \\ 1 - 7z + 5t & -1 \end{vmatrix} = -6 + 26z - 17t$$

De esta manera,

$$x = \frac{D_x}{D} = \frac{-6 + 26z - 17t}{-1}$$

y por lo tanto, las soluciones se escriben

$$\begin{cases} x = 6 - 26 + 17t \\ y = -1 + 7z - 5t \\ z = z \\ t = t \end{cases}$$

para argumentos reales arbitrarios de z y t.

Este último sistema define a un plano de dimensión 2 contenido en $\mathbb{R}^4$, esto es, la intersección de los hiperplanos dados de dimensión 3 es un plano de dimensión 2. $\triangleright$

Distancia de un punto a un plano

Dados, un plano P y un punto $p \in \mathbb{R}^3$, la distancia de p al plano P es la distancia ortogonal de p a los puntos q del plano P.

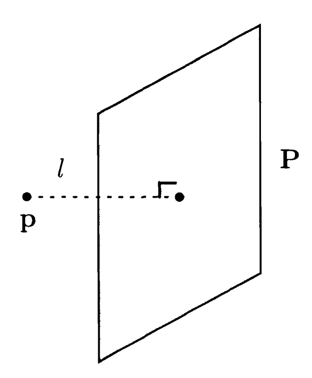


Figura 2.16: Distancia de un punto a un plano.

Para calcular la distancia mencionada, se procede de la siguiente manera.

$\triangleleft$ Si el plano P se define por la ecuación

$$ax + by + cz = d$$

entonces $n = (a, b, c)$ es un vector normal p.

Por otro lado, cuando el punto $q \in P$ nos da la distancia ortogonal buscada, entonces el vector $p - q$ es múltiplo de n, es decir,

$$p - q = tn$$

para algún escalar t (ya que son paralelos), lo que implica que

$$||p - q|| = |t|\,||n||.$$

Por otro lado, al buscar $q \in P$, este punto está en la recta con un vector n que se une a p con q, escrita de la forma

$$\begin{cases} x = t\,a + p_1 \\ y = t\,b + p_2 \\ z = t\,c + p_3 \end{cases}$$

donde $p = (p_1, p_2, p_3)$ y $n = (a, b, c)$ es el vector director de la recta.

Buscamos entonces un escalar t tal que la recta contenga un punto $q \in P$ q satisfaciendo el sitema que define a la recta, pero además q esté también contenido en P. Para esto, $q = (x, y, z) \in L$ debe también satisfacer la ecuación del plano, es decir,

$$d = a(ta + p_1) + b(tb + p_2) + c(tc + p_3) = t(a^2 + b^2 + c^2) + ap_1 + bp_2 + cp_3$$
$$= (ap_1 + bp_2 + cp_3) + t(a^2 + b^2 + c^2)$$

lo que indica, al despejar t,

$$t = \frac{d - (ap_1 + bp_2 + cp_3)}{(a^2 + b^2 + c^2)} = \frac{d - (ap_1 + bp_2 + cp_3)}{||n||^2}$$

Por otro lado, el punto $q \in P$ debe satisfacer además la relación

$$||q - p|| = |t| \, ||n||$$

siendo $||q - p||$ la distancia buscada de q a p.

Si entendemos por ℓ a tal **distancia**, se tiene

$$\ell = |t| \, ||n|| = \frac{|d - (ap_1 + bp_2 + cp_3)|}{||n||^2} ||n||$$

lo que implica

$$\ell = \frac{|d - (ap_1 + bp_2 + cp_3)|}{||n||}$$

Se resume la anterior discusión en el siguiente lema.

LEMA 2.10 *La distancia ℓ del punto $p = (p_1, p_2, p_3)$ al plano P definido por la ecuación*

$$ax + by + cz = d$$

se calcula por la fórmula

$$\ell = \frac{|d - (ap_1 + bp_2 + cp_3)|}{||n||}$$

donde $n = (a, b, c)$ es el normal a P

Si $p = (p_1, p_2, p_3)$ está en el plano, entonces satisface la ecuación del plano, es decir,

$$0 = d - (ap_1 + bp_2 + cp_3)$$

y, por lo tanto, su distancia al plano P es cero.



EJEMPLO. Sea el plano P dado por la ecuación

$$3x + 2y - z = -1$$

y tomemos el punto $p = (2,5,6)$.

$\triangleleft$ La distancia ℓ del punto p al plano P es calculada mediante la cadena de igualdades,

$$\ell = \frac{|-1-[3(2)+2(5)-6]|}{\sqrt{9+4+1}} = \frac{|-1-10|}{\sqrt{14}} = \frac{11}{\sqrt{14}} \quad \triangleright$$

Ejercicios

1. Calcular $< \xi, \eta >$ para cada pareja de vectores dada.

a. $\xi = (-1,2), \quad \eta = (3,1)$
b. $\xi = (-1,0,1), \quad \eta = (1,2,3)$
c. $\xi = (-1,2,3), \quad \eta = (1,2,-4)$
d. $\xi = (1,1,0,-1), \quad \eta = (-1,0,1,2)$
e. $\xi = (4,-2,3,-1), \quad \eta = (-1,-2,3,-4)$
¿Cuáles de las parejas de vectores dados son ortogonales?

2. Calcular $||\xi||$ y $||\eta||$ en cada uno de los incisos del ejercicio **1.**

3. Calcular el coseno del ángulo formado por cada pareja de vectores ξ y η del ejercicio **1.**

4. Calcular los ángulos del triángulo formado por los puntos
a. $(-1,1,1), \quad (1,-2,3), \quad (3,-1,1)$
b. $(1,0,0), \quad (0,1,0), \quad (0,0,1)$

5. **a.** Calcular la proyección de ξ y η de todos los vectores del ejercicio **1.**
b. Calcular la proyección de η en ξ de todos los vectores del ejercicio **1.**

6. Demostrar que para dos vectores ξ, η con ángulo θ entre ellos se cumple la fórmula

$$||\xi - \eta||^2 = ||\xi||^2 + ||\eta||^2 - 2||\xi||||\eta|| \cos\theta$$

7. Calcular el producto vectorial $[\xi, \eta]$ de los vectores
a. $\xi = (2,3,-1) \quad \eta = (6,2,3)$
b. $\xi = (-6,0,5) \quad \eta = (3,0,3)$
c. $\xi = (1,0,-1) \quad \eta = (-1,0,2)$
d. $\xi = (1,2,3) \quad \eta = (2,4,6)$
¿Cuáles de las parejas de vectores dados son paralelos?

8. Calcular el área del paralelogramo formado por los vectores ξ y η en cada inciso del ejercicio **7.**



9. Calcule el volumen del paralelepípedo formado por las siguientes ternas de vectores. Calcule primero su triple producto escalar.
a. $\xi = (2,0,1), \quad \eta = (1,3,2), \quad \zeta = (4,-1,1)$
b. $\xi = (-2,1,1), \quad \eta = (-4,3,2), \quad \zeta = (0,0,1)$
c. $\xi = (0,-1,-1), \quad \eta = (0,0,-1), \quad \zeta = (-1,2,3)$
d. ¿Cuáles son una base de vectores en $\mathbb{R}^3$?
e. Escriba el vector $(-3,-4,1)$ como combinación lineal de cada base encontrada.

10. ¿Cuáles de las siguientes matrices son diagonalizables?

a. $A = \begin{pmatrix} -11 & 12 \\ -20 & 17 \end{pmatrix}$ b. $A = \begin{pmatrix} -2 & 3 \\ -1 & 1 \end{pmatrix}$

c. $A = \begin{pmatrix} 1 & 0 \\ 1/2 & 1 \end{pmatrix}$ d. $A = \begin{pmatrix} 1 & 0 \\ 6 & -1 \end{pmatrix}$

e. $A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ **f.** $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix}$

g. $A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & -3 & 0 \\ 4 & -13 & 1 \end{pmatrix}$ **h.** $A = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

i. $A = \begin{pmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{pmatrix}$ j. $A = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix}$

Para el caso afirmativo, calcule la matriz B, que cambia la base canónica por la base de vectores propios, tal que $B^{-1}AB$ sea diagonal. Compruebe, realizando el cálculo, que $B^{-1}AB$ es diagonal, y que los vectores propios que conforman B son una base.

11. Encontrar la ecuación paramétrica de la recta que pasa por los puntos indicados

a. $(1,1,-1)$ y $(4,3,2)$
b. $(0,1,-6)$ y $(-4,0,2)$
c. $(1,-2)$ y $(2,1)$
d. $(1,1,-1,1)$ y $(4,4,-4,4)$

12. Calcular la ecuación paramétrica de la recta por p con dirección ξ para cada pareja dada.

a. $p = (1,1,-1), \quad \xi = (-2,3,-1)$
b. $p = (-2,1), \quad \xi = (4,2)$
c. $p = (2,3,0), \quad \xi = (-1,2,3)$



d. $p = (0,1,1,2)$, $\xi = (-1,3,4,2)$

13. Dar la ecuación de la recta que tiene un vector director ξ que es ortogonal a los vectores $\zeta = (1,1,1)$ y $\eta = (-1,0,1)$, y que pasa por el punto $(-3,1,2)$.

14. Calcular la ecuación del plano que pasa por el punto dado p y tiene vector normal n para cada pareja dada.

a. $n = (-1,1,2)$, $p = (2,3,-1)$

b. $n = (4,-2,3)$, $p = (-1,4,1)$

c. $n = (-1,5,3)$, $p = (4,1,6)$

15. Calcular la ecuación del plano que pasa por los tres puntos dados

a. $(1,1,1)$, $(2,1,-1)$, $(3,0,1)$

b. $(3,-2,-1)$, $(1,2,3)$, $(2,-3,1)$

c. $(4,0,2)$, $(2,-1,0)$, $(0,2,3)$

16. Calcular el ángulo con el que se intersectan los planos

a. definidos por **a.** y **b.** en el ejercicio **14.**

b. definidos por **a.** y **b.** en el ejercicio **15.**

c. definidos por **c.** del ejercicio **14** y **c.** del ejercicio **15.**

d. $\begin{cases} x+y+x=1 \\ -x+y-z=0 \end{cases}$

e. $\begin{cases} 2x-y-z=3 \\ 3x+y-2z=1 \end{cases}$

17. Calcular la intersección de las parejas de planos dadas en todos los incisos del ejercicio **16.**

18. Calcular la ecuación de la recta que pasa por el punto p y cuyo vector director ξ es el normal al plano P cuando:

a. $p = (1,1,1)$ y P pasa por los puntos $(1,1,-1), (-2,3,1), (6,2,0)$

b. $p = (1,0,3)$ y P es ortogonal a los planos

$$\begin{cases} x-y+z=2 \\ -x+y-z=1 \end{cases}$$

c. $p = (2,3,1)$ y P es paralelo al plano qua pasa por los puntos $(4,2,-1)$, $(0,1,-3), (1,0,2)$

19. Calcular la intersección de la recta L con el plano P, si

a. L y P son la recta y el plano de **a.** en el ejercicio 18.

b. L y P son la recta y el plano de **c.** en el ejercicio **18.**

c. L es la recta por $p = (1,3,4)$ con vector director $\xi = (2,1,0)$, y P es el plano con ecuación $2x - y + z = 4$



20. Calcular la distancia del punto p al plano P si

a. $p = (1, 0, 3)$, $P : 6x - 8y + z = 6$.

b. $p = (-2, 1, 3)$, $P : x + y + z = 2$.

c. $p = (4, 3, 2)$, P pasa por los puntos. $(1, 1, 1), (0, 0, 0), (-1, 1, -1)$


Casa abierta al tiempo

Capítulo 3

Curvas en $\mathbb{R}^2$ y en $\mathbb{R}^3$

3.1 Curvas suaves

En este capítulo estudiamos aquellos objetos geométricos en $\mathbb{R}^3$ conocidos como curvas. Lo hacemos de dos maneras, como la imagen de una función vectorial de variable real, y como objetos para los cuales localmente se ven como una imagen de cierta función vectorial de variable real (curvas parametrizadas). Iniciamos la discusión con un ejemplo.

EJEMPLO. Una recta en $\mathbb{R}^3$ es un objeto que se describe por un punto donde pasa $p = (p_1, p_2, p_3)$ y un vector director $\xi = (\xi^1, \xi^2, \xi^3)$.

$\lhd$ Su ecuación se describe como una relación de tipo

$$\gamma : \mathbb{R} \to \mathbb{R}^3$$

donde a cada punto $t \in \mathbb{R}$ se asocia un vector formado por coordenadas que son funciones reales de variable real

$$\gamma(t) = (x(t),\ y(t),\ z(t)),$$

siendo las funciones coordenadas definidas por el argumento escalar mencionado,

$$x = x(t) = \xi^1 t + p_1$$
$$y = y(t) = \xi^2 t + p_2$$
$$z = z(t) = \xi^3 t + p_3 \quad \rhd$$

DEFINICIÓN. Una **función vectorial** de argumento escalar es una relación

$$\gamma : (a, b) \to \mathbb{R}^3$$



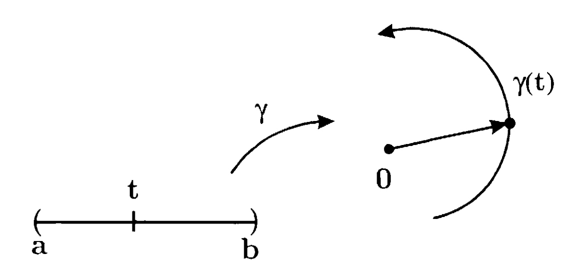


Figura 3.1: Función vectorial de variable real.

tal que a cada $t \in (a, b)$ se le asocia un vector $\gamma(t)$ en $\mathbb{R}^3$.

En coordenadas x, y, z de $\mathbb{R}^3$, tal asociación se escribe

$$\gamma(t) = (x(t),\ y(t),\ z(t))$$

donde $x, y, z : (a, b) \to \mathbb{R}$ son funciones reales (definidas sobre $\mathbb{R}$) de variable real t. Esto es, las funciones coordenadas se escriben

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases} , \quad t \in (a, b)$$

En otra notación, utilizando la base canónica $\{e_1, e_2, e_3\}$ de $\mathbb{R}^3$, podemos escribir tal asociación mediante la igualdad

$$\gamma(t) = x(t)e_1 + y(t)e_2 + z(t)e_3$$

A lo largo de este trabajo, utilizaremos preferentemente la notación $\gamma(t) = (x(t), y(t), z(t))$ que no involucra a los vectores canónicos.

EJEMPLO. Sea la función $\gamma : \mathbb{R} \to \mathbb{R}^2$ dada por la relación

$$\gamma(t) = (\cos t, \operatorname{sen} t)$$

$\lhd$ Es una función vectorial (a $\mathbb{R}^2$) con argumento escalar dado por la variable $t \in \mathbb{R}$.

Hacemos una asignación de argumentos y calculamos sus valores, como lo muestra la siguiente tabla.

t	$\gamma(t)$
0	$(1,0)$
$\pi/4$	$\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$
$\pi/2$	$(0,1)$
$\frac{3\pi}{4}$	$\left(-\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$
π	$(-1,0)$
$(5/4)\pi$	$\left(-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)$
$(3/2)\pi$	$(0,-1)$
$(7/4)\pi$	$\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)$
2π	$(1,0)$
$(9/4)\pi$	$\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$

A continuación mostramos cómo se consiguieron algunos elementos de la tabla.

$$\gamma(0) = (\cos 0, \operatorname{sen} 0) = (1,0)$$

$$\gamma(\pi/4) = (\cos \pi/4, \operatorname{sen} \pi/4) = \left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$$

$$\gamma(\pi/2) = (\cos \pi/2, \operatorname{sen} \pi/2) = (0,1)$$

$$\gamma\left(\frac{3\pi}{4}\right) = \left(-\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$$

$$\gamma(\pi) = (-1,0)$$

$$\gamma((5/4)\pi) = \left(-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)$$

$$\gamma((3/2)\pi) = (0,-1)$$

Observamos que todos los valores de γ obtenidos de esta manera están contenidos en el círculo unitario del plano $\mathbb{R}^2_{x,y}$ como se muestra en la figura 3.2.

Afirmamos que la imagen de $\mathbb{R}$ bajo γ es completamente el círculo unitario

$$S^1 = \{(x,y)|\, x^2+y^2=1\}$$

Para demostrar esto, sea $\gamma(t)$ un punto en la imagen, entonces,

$$||\gamma(t)||^2 = ||(\cos t,\, \operatorname{sen} t)||^2 = \cos^2 t + \operatorname{sen}^2 t = 1$$



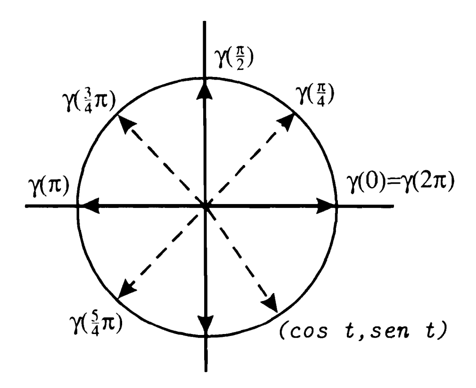


Figura 3.2: Círculo unitario de S^1

por lo tanto, ya que $||\gamma(t)|| = 1$ entonces, para cualquier $t \in \mathbb{R}$, $\gamma(t) \in S^1$

Ahora sea $(x, y) \in S^1$ y sea t el ángulo que forma el vector (x, y) con el eje x. Si sobre el círculo se localizan las coordenadas (x, y) de un punto. En virtud de que el radio de S^1 es 1, se tienen las igualdades trigonométricas

$$\begin{cases} x = \cos t \\ y = \operatorname{sen} t \end{cases}$$

de donde $(x, y) \in S^1$ está en la imagen γ.

De esta manera, la imagen de γ es un **círculo unitario**.

Como las funciones $\operatorname{sen} t$, $\cos t$ son 2π- periódicas, entonces γ enrrolla al círculo S^1 en contra de las manecillas del reloj un número infinito (numerable) de veces. $\triangleright$

EJEMPLO. Consideremos la función vectorial $\gamma : (-\infty, \infty) \to \mathbb{R}^2$ dada por la regla de correspondencia

$$\gamma(t) = (R\cos t,\ R\ \operatorname{sen} t)$$

donde R es un número real positivo.

$\triangleleft$ Tendremos en este caso que, al escribir tal relación como

$$\gamma(t) = (R\cos t,\ R\ \operatorname{sen} t) = R(\cos t,\ \operatorname{sen} t)$$

ésta es una deformación del círculo unitario a un círculo de radio $R > 0$ con centro en el origen de coordenadas (véase la figura 3.3). $\triangleright$

EJEMPLO. Sea ahora la función $\gamma : (-\infty, \infty) \to \mathbb{R}^2$ dada por la relación

$$\gamma(t) = (4\cos t,\ 4\operatorname{sen} t,\ 2t)$$

$\triangleleft$ γ es la función que a cada t le asocia un punto en $\mathbb{R}^3$, mediante las relaciones

$$\begin{cases} x(t) = 4\cos t \\ y(t) = 4\operatorname{sen} t \\ z(t) = 2t \end{cases}$$

Al olvidar momentáneamente la variable z, se tiene que los puntos con coordenadas (x, y) en la imagen de γ se mueven en un círculo de radio 4, según el ejemplo anterior. La variable z sólo se mueve con su dirección, dirigiéndose positivamente.

Esto nos dice que la imagen $\gamma(t)$ está contenida en la superficie dada por $x^2 + y^2 = 16$ que es un cilindro circular de radio 4 con directriz del eje z.

Al evaluar γ en los extremos del intervalo $[0, 2\pi]$ se obtienen los puntos

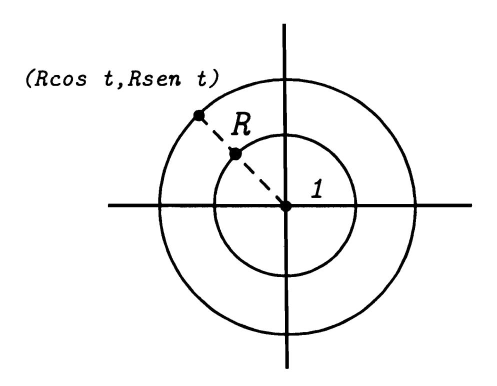


Figura 3.3: Círculo de radio R.

$$\gamma(0) = (4, 0, 0)$$

$$\gamma(2\pi) = (4, 0, 4\pi)$$

que quedan verticalmente situados en dirección del eje z, y a una distancia (vertical) igual al 4π, lo que nos dice que la imagen de γ es una **hélice** de radio 4 y con paso de altura 4π (que se repite a intervalos de t iguales a 2π) como lo muestra la figura 3.4. $\triangleright$

Iniciamos ahora la discusión de la diferenciabilidad de una curva espacial, utilizando para ello los conceptos y resultados clásicos del cálculo diferencial en una variable (véase **Reyes**, 1996).

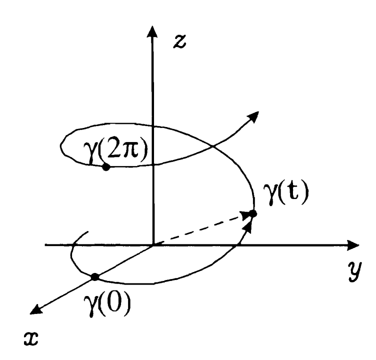


Figura 3.4: Hélice de radio 4 con paso de altura 4π.

Sea $\gamma : (a, b) \to \mathbb{R}^3$ una función vectorial de variable real definida en el intervalo $(a, b) \subset \mathbb{R}$.

Sean $t_0, t \in (a, b)$ con el argumento temporal variable t próximo del tiempo prefijado t_0.

Calculamos la diferencia entre sus vectores imágenes $\gamma(t) - \gamma(t_0)$ y luego la comparamos vía un cociente con la diferencia de argumentos (lo cual es posible pues podemos dividir vectores entre escalares).

$$\frac{\gamma(t) - \gamma(t_0)}{t - t_0} \left(= \frac{\text{diferencia de imágenes}}{\text{diferencia temporal}} \right)$$

Si pensamos en las imágenes como posiciones de un cuerpo puntual que se mueve en el espacio, tal cociente representaría un **vector velocidad promedio** en el intervalo $[t_0, t]$, y el proceso de límite nos daría un **vector velocidad** instantánea (en caso de existir) al tiempo t_0.

Definimos tal proceso de límite en el espacio, mediante los límites de las funciones coordenadas que son ya de conocimiento del cálculo diferencial de funciones reales de variable real.

DEFINICIÓN. Se dice que la función $\gamma : (a, b) \to \mathbb{R}^3$ es **diferenciable** en el punto $t_0 \in (a, b)$, si existe el límite vectorial

$$\lim_{t \to t_0} \frac{\gamma(t) - \gamma(t_0)}{t - t_0} = \lim_{t \to t_0} \frac{(x(t), y(t), z(t)) - (x(t_0), y(t_0), z(t_0))}{t - t_0}$$

$$= \lim_{t \to t_0} \left(\frac{x(t) - x(t_0)}{t - t_0}, \frac{y(t) - y(t_0)}{t - t_0}, \frac{z(t) - z(t_0)}{t - t_0} \right)$$



En tal caso, tal vector se definirá por

$$\frac{d\gamma}{dt}(t_0) = \lim_{t\to t_0} \frac{\gamma(t) - \gamma(t_0)}{t - t_0}$$

y se llamará la **derivada** de γ en el punto t_0.

Si en cada punto del intervalo la función γ es diferenciable, se dirá simplemente que es una **función diferenciable**.

Ya que el límite que define a $\frac{d\gamma}{dt}(t_0)$ depende de la existencia de cada uno de los límites por coordenadas (según la última igualdad en la definición), entonces tenemos,

LEMA 3.1 *La función γ es diferenciable en t_0, si y sólo sí, cada función coordenada es diferenciable en t_0. Esto es, si cada límite*

$$\frac{dx}{dt}(t_0) = \lim_{t\to t_0} \frac{x(t) - x(t_0)}{t - t_0},$$

$$\frac{dy}{dt}(t_0) = \lim_{t\to t_0} \frac{y(t) - y(t_0)}{t - t_0},$$

$$\frac{dz}{dt}(t_0) = \lim_{t\to t_0} \frac{z(t) - z(t_0)}{t - t_0},$$

existe. De aquí que γ es diferenciable, si y sólo sí, cada una de sus funciones coordenadas es diferenciable.

De esta manera, el problema de diferenciabilidad de una curva se ha reducido a la diferenciabilidad de las funciones coordenadas. Por eso podemos iniciar una discución de una propiedad más fuerte de las funciones vectoriales de variable real, es decir, de las **curvas diferenciables**. [1]

Por una **función vectorial suave de clase** C^r entenderemos a una función vectorial cuyas funciones coordenadas son diferenciables de clase C^r con $r \geq 1$.

De esta manera, γ es una función suave, si cada una de sus funciones coordenadas es una función real de variable real suave en el intervalo $J = (a, b)$.

Por El **teorema de Whitney** (Teorema 4.9 de **Reyes**, 1996), cada una de las funciones coordenadas de una función vectorial de variable real cuando es continua, puede aproximarse por una función de clase C^r. Así, en las aplicaciones, cuando se piensa en una función vectorial de variable

[1] El lector no familiarizado con la clase de diferenciabilidad de una función real puede consultar el capítulo 3 de Reyes, 1996.

real, ésta se supone de una clase de diferenciabilidad lo suficientemente alta, de acuerdo a las necesidades del problema que se esté solucionando. Esto es, en las aplicaciones, las curvas espaciales genéricamente se piensan suaves.

Si la curva γ es suave, definimos a su derivada en el punto arbitrario $t \in J = (a, b)$ usando una notación clásica:

$$\frac{d\gamma}{dt} = \dot{\gamma}(t) = (\dot{x}(t), \dot{y}(t), \dot{z}(t))$$

donde el punto sobre la variable define a la derivada respecto al tiempo t. El vector derivada $\dot{\gamma}(t)$ es un vector que además de ser tangente a la curva γ en el punto $\gamma(t)$, es un vector localizado posicionado en el punto $\gamma(t)$.

DEFINICIÓN. Una función vectorial $\gamma : (a, b) \to \mathbb{R}^3$ se llamará una **curva regular**, si para todo $t \in (a, b)$ su derivada $\dot{\gamma}(t)$ no se anula.

EJEMPLO. Sea la función vectorial

$$\gamma(t) = (\cos t, \operatorname{sen} t, t),$$

entonces la curva de imagen γ en el intervalo $(a, b) = (-\infty, \infty)$ es una hélice de paso de altura 2π, contenida en el cilindro $x^2 + y^2 = 1$ del espacio $\mathbb{R}^3$.

Claramente

$$\dot{\gamma}(t) = (-\operatorname{sen} t, \cos t, 1)$$

es un vector que nunca se anula y, por ello, es que la curva es regular. (véase la figura 3.5 a.).

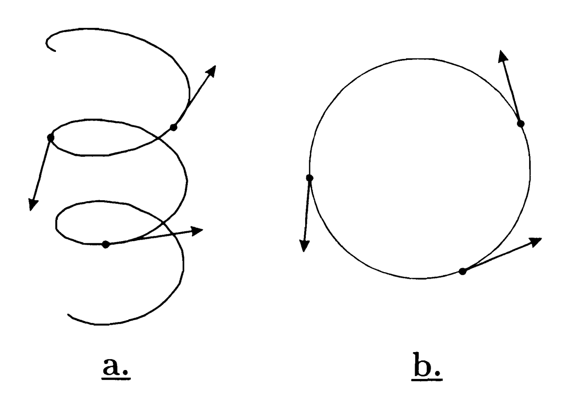


Figura 3.5: Curvas imágenes de funciones vectoriales

EJEMPLO. Sea la función vectorial $\gamma : (-\infty, \infty) \to \mathbb{R}^2$ dada por

$$\gamma(t) = (R\cos t, R\,\mathrm{sen}\, t).$$

Entonces su imagen es un círculo de radio $R > 0$ con centro en el origen del plano $\mathbb{R}^2$.

Su derivada $\dot{\gamma}(t) = (-R\,\mathrm{sen}\, t, R\cos t)$ es un vector que no se anula y, por lo tanto, se entenderá también como una **curva regular** (véase la figura 3.5 b.).

EJEMPLO. Considérese la curva $\gamma : J \to \mathbb{R}^3$ dada por

$$\gamma(t) = (t^3 - t, t^2 - 1)$$

$\lhd$ Es claro que el dominio de la curva γ es todo el eje real $J = (-\infty, \infty)$.

Inicialmente, para hacer el dibujo de la imagen de esta curva, asignamos algunos argumentos a la función y calculamos sus valores

t	-3	-2	-1	0	1	2	3
$\gamma(t)$	$(-24, 8)$	$(-6, 3)$	$(0, 0)$	$(0, -1)$	$(0, 0)$	$(6, 3)$	$(24, 8)$

Teniendo estos valores vectoriales en el plano, utilizamos ahora un método analítico para dibujar más precisamente la imagen de γ.

Procedemos a analizar los vectores imágenes de la curva plana según los argumentos en el dominio, utilizando para ello las funciones coordenadas.

A fin de resolver la desigualdad $t^3 - t \geq 0$ para la primer función coordenada, calculamos las raíces de $t^3 - t = 0$ que son $t = 0, 1, -1$.

Estos números nos dividen el intervalo $(-\infty, \infty)$ en tres partes que nos permiten resolver tal desigualdad, sabiendo que

$$t^3 - t \geq 0 \iff t \in [-1, 0] \quad \text{ó} \quad t \in [1, \infty].$$

Consecuentemente, para la primer función coordenada se tiene

$$t^3 - t \geq 0 \iff t \in [-1, 0] \cup [1, \infty)$$

$$t^3 - t < 0 \iff t \in (-\infty, -1) \cup (0, 1)$$

Análogamente para la segunda función coordenada, la ecuación $t^2 - 1 = 0$, tiene las raíces $t = \pm 1$, lo que implica,

$$t^2 - 1 \geq 0 \iff (-\infty, -1] \cup [1, \infty)$$

$$t^2 - 1 < 0 \Longleftrightarrow (-1, 1)$$

Evaluamos a γ en los puntos importantes $t = 0, -1, 1$ obtenidos al resolver las desigualdades anteriores:

$$\gamma(-1) = ((-1)^3 - (-1), (1)^2 - 1) = (0, 0)$$

$$\gamma(0) = ((0)^3 - (0), 0^2 - 1) = (0, -1)$$

$$\gamma(1) = (-1^3 - 1, 1^2 - 1) = (0, 0)$$

y al dividir al dominio J en los subintervalos

$$(-\infty, -1),\ (-1, 0),\ (0, 1), (1, \infty)$$

se obtiene la tabla

	$(-\infty, -1)$	$(-1, 0)$	$(0, -1)$	$(1, -\infty)$
$x(t)$	$-$	$+$	$-$	$+$
$y(t)$	$+$	$-$	$-$	$+$

que indica la posición de la imagen de un argumento sobre la curva γ en los cuadrantes del plano (véase figura 3.6), según las signaturas de sus coordenadas.

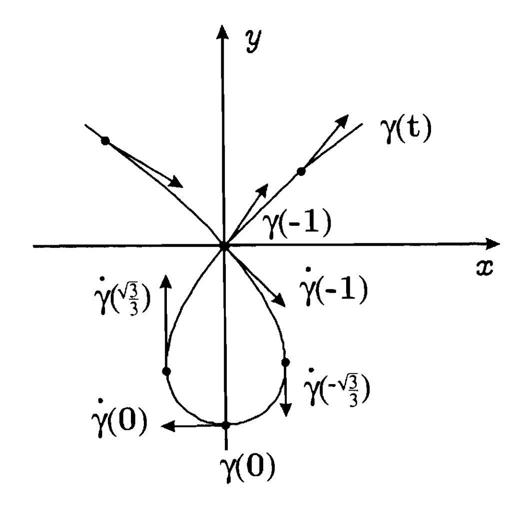


Figura 3.6: Posición de la imagen de la curva $\gamma(t) = (t^3 - t, t^2 - 1)$

Ahora realizamos un análisis cualitativo de γ, usando su derivada

$$\dot{\gamma}(t) = (3t^2 - 1, 2t)$$

mediante un método similar al usado para la signatura de γ en el plano.

Para la primer coordenada $\dot{x}(t)$ de $\dot{\gamma}(t)$ se tiene

$$\dot{x}(t) = 3t^2 - 1 > 0 \Longleftrightarrow \quad 3t^2 - 1 > 0 \Longleftrightarrow 3t^2 > 1$$

de donde

$$\dot{x}(t) > 0 \Longleftrightarrow t \in \left(-\infty, \frac{\sqrt{3}}{3}\right) \cup \left(\frac{\sqrt{3}}{3}, \infty\right)$$

Por lo tanto,

$$\dot{x}(t) = 3t^2 - 1 < 0 \Longleftrightarrow t \in \left(-\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right)$$

Análogamente, para la segunda coordenada de $\dot{\gamma}(t)$,

$$\dot{y}(t) = 2t > 0 \Longleftrightarrow x \in (0, \infty)$$

$$\dot{y}(t) = 2t < 0 \Longleftrightarrow x \in (-\infty, 0)$$

De las dos desigualdades anteriores se consiguen los puntos importantes $t = 0, -\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}$.

Ahora calculamos la imágenes bajo γ y $\dot{\gamma}$ de estos puntos importantes:

$$\gamma(0) = (0, -1)$$

$$\gamma\left(\frac{-\sqrt{3}}{3}\right) = \left(\frac{2\sqrt{3}}{9}, -\frac{2}{3}\right)$$

$$\gamma\left(\frac{\sqrt{3}}{3}\right) = \left(\frac{-2\sqrt{3}}{9}, -\frac{2}{3}\right)$$

$$\dot{\gamma}(0) = (-1, 0)$$

$$\dot{\gamma}\left(-\frac{\sqrt{3}}{3}\right) = \left(0, -\frac{2\sqrt{3}}{3}\right)$$

$$\dot{\gamma}\left(\frac{\sqrt{3}}{3}\right) = \left(0, \frac{2\sqrt{3}}{3}\right)$$

La figura 3.6 ilustra a tales puntos y a sus vectores velocidad (tangentes) correspondientes.

Con los intervalos

$$\left(-\infty, -\frac{\sqrt{3}}{3}\right), \left(-\frac{\sqrt{3}}{3}, 0\right), \left(0, \frac{\sqrt{3}}{3}\right), \left(\frac{\sqrt{3}}{3}, \infty\right)$$

obtenidos mediante este segundo análisis se forma la siguiente tabla que describe la signatura de las coordenadas de $\dot{\gamma}(t) = (\dot{x}), \dot{y}(t))$.

	$\left(-\infty, -\frac{\sqrt{3}}{3}\right)$	$\left(-\frac{\sqrt{3}}{3}, 0\right)$	$\left(0, \frac{\sqrt{3}}{3}\right)$	$\left(\frac{\sqrt{3}}{3}, \infty\right)$
$\dot{x}(t)$	+	−	−	+
$\dot{y}(t)$	−	−	+	+
$\dot{\gamma}(t)$	↘	↙	↖	↗

La flecha del renglón correspondiente a $\dot{\gamma}(t)$ en cada subintervalo describe cualitativamente la posición del vector velocidad tangente a $\gamma(t)$ (véase la figura 3.6).

Además, observamos que la imagen de γ tiene una autointersección en $\gamma(1) = \gamma(-1)$ lo que le hace ser una función que no es inyectiva. En el punto $\gamma(1) = \gamma(-1) = (0, 0)$ se tienen dos vectores derivadas

$$\dot{\gamma}(1) = (3t^2 - 1, 2t)|_1 = (2, 2)$$

$$\dot{\gamma}(-1) = (3t^2 - 1, 2t)|_{-1} = (2, -2)$$

De esta forma un trazo de la curva en una región acotada alrededor de $(0, 0)$ estaría dado por la figura 3.6.

Hacemos el análisis a los extremos del dominio $J = (-\infty, \infty)$.

Si $t \to -\infty$, entonces,

$$\lim_{t\to-\infty} \gamma(t) = \lim_{t\to-\infty} (t^3 - t, t^2 - 1) = (\lim_{t\to-\infty} (t^3 - t), \lim_{t\to-\infty} (t^2 - 1)) \to (-\infty, \infty)$$

por El teorema de alto orden (véase Teorema 2.23 de Reyes, 1996).

Análogamente, si $t \to +\infty$, entonces, por el mismo teorema de alto orden, se tiene

$$\lim_{t\to+\infty} \gamma(t) = (\lim_{t\to+\infty} t^3, \lim_{t\to+\infty} t^2) \to (\infty, \infty)$$

lo que se puede observar en la figura 3.6. ▷

A los vectores derivadas $\dot{\gamma}(t)$ se les entenderá también como **vectores velocidad** de γ en el punto $\gamma(t)$. A la longitud de $\dot{\gamma}(t)$, es decir a $||\dot{\gamma}(t)||$ se le llamará la **rapidez** de γ en el punto $\gamma(t)$.

EJEMPLO. (La cúspide) Sea la función vectorial dada por

$$\gamma(t) = (t^3, t^2)$$

$\lhd$ Es claro que el dominio de la curva γ es $J = (-\infty, \infty)$.

Asignamos primeramente algunos argumentos a la función γ calculando sus valores, y obteniendo la siguiente tabla.

t	-3	-2	-1	0	1	2	3
$\gamma(t)$	$(-27,9)$	$(-8,4)$	$(-1,1)$	$(0,0)$	$(1,1)$	$(8,4)$	$(27,9)$

Procedemos de forma análoga al ejemplo anterior, estudiando la signaturas de la función y su derivada.

Para la primer función coordenada, resolvemos la desigualdad $x(t) = t^3 \geq 0$ que simplemente es equivalente a $t \geq 0$.

Por lo tanto,

$$x(t) = t^3 \geq 0 \Longleftrightarrow t \in [0, \infty]$$

y consecuentemente,

$$x(t) = t^3 < 0 \Longleftrightarrow t \in [-\infty, 0)$$

Análogamente para la segunda coordenada, obtenemos,

$$y(t) = t^2 \geq 0 \Longleftrightarrow t \in (-\infty, \infty)$$

Evaluamos los puntos importantes encontrados, que es apenas $t = 0$,

$$\gamma(0) = (0,0)$$

y con los subintervalos obtenidos se tiene la tabla de signatura de $\gamma(t)$,

	$(-\infty,0)$	$(0,\infty)$
$x(t)$	$-$	$+$
$y(t)$	$+$	$+$

Analizamos la signatura de la derivada de la curva,

$$\gamma(t) = (3t^2, 2t)$$

resolviendo primeramente la desigualdad,

$$\dot{x}(t) = 3t^2 > 0 \Longleftrightarrow t \in (-\infty, 0) \cup (0, \infty)$$

Por otro lado,

$$\dot{y} = 2t > 0 \Longleftrightarrow t \in (0, \infty)$$
$$\dot{y}(t) = 2t < 0 \Longleftrightarrow t \in (-\infty, 0)$$

y se tiene en este caso, nuevamente el punto importante $t = 0$.

Calculemos las imágenes bajo γ y $\dot{\gamma}$ de este punto importante

$$\gamma(0) = (0, 0)$$
$$\dot{\gamma}(0) = (0, 0)$$

que hace de $\gamma(t)$ una curva irregular.

Ahora con los intervalos $(-\infty, 0)$, $(0, \infty)$ se forma la siguiente tabla de signaturas de las componentes de los vectores velocidad,

	$(-\infty, 0)$	$(0, \infty)$
$\dot{x}(t)$	+	+
$\dot{y}(t)$	−	+
$\dot{\gamma}(t)$	↘	↗

Finalmente, hacemos el análisis a los tiempos extremos del dominio $J = (-\infty, \infty)$, calculamos los límites,

$$\lim_{t \to -\infty} \gamma(t) = \lim_{t \to -\infty} (t^3, t^2) = (\lim_{t \to -\infty} t^3, \lim_{t \to -\infty} t^2) \to (-\infty, \infty)$$

y

$$\lim_{t \to \infty} \gamma(t) = \lim_{t \to \infty} (t^3, t^2) = (\lim_{t \to \infty} t^3, \lim_{t \to \infty} t^2) \to (\infty, \infty)$$

La figura 3.7 ilustra el trazo de la curva imagen de esta función llamada **cúspide**. ▷

Damos ahora unas propiedades aritméticas de la derivada de funciones vectoriales con argumento escalar.

LEMA 3.2 *Sean $\gamma_1, \gamma_2 : (a, b) \to \mathbb{R}^3$ dos funciones vectoriales suaves de variable real.*

a. *Para los escalares λ, $\mu \in \mathbb{R}$, se verifica fácilmente que*

$$\frac{d}{dt}(\lambda\gamma_1 + \mu\gamma_2)(t) = \lambda\frac{d\gamma_1}{dt}(t) + \mu\frac{d\gamma_2}{dt}(t) = \lambda\dot{\gamma_1}(t) + \mu\dot{\gamma_2}(t)$$

b. *Para una función real de variable real suave $\lambda = \lambda(t)$ definida en el intervalo $(a, b) \subset \mathbb{R}$ es justa la relación*

$$\frac{d}{dt}(\lambda(t)\gamma(t)) = \frac{d\lambda}{dt}(t)\gamma(t) + \lambda(t)\frac{d\gamma}{dt}(t) = \dot{\lambda}(t)\gamma(t) + \lambda(t)\dot{\gamma}(t)$$



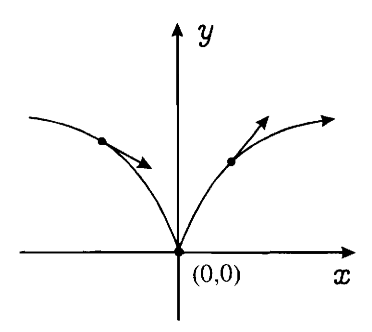


Figura 3.7: Curva imagen de la cúspide.

◁ Son cálculos directos y omitimos su prueba ▷.

Sean ahora dos funciones vectoriales suaves $\gamma_1, \gamma_2 : (a,b) \to \mathbb{R}^3$ en todo el intervalo, y consideremos para cada $t \in (a,b)$ los vectores velocidad $\dot{\gamma}_1(t)$, $\dot{\gamma}_2(t)$. Se tiene el siguiente resultado acerca del cálculo de la derivada en productos escalar y vectorial.

LEMA 3.3 *Si $\gamma_1, \gamma_2 : (a,b) \to \mathbb{R}^3$ son funciones suaves, son justas las fórmulas de Leibniz:*

a.

$$\frac{d}{dt} < \gamma_1(t), \gamma_2(t) >= \left\langle \frac{d\gamma_1}{dt}(t),\, \gamma_2(t) \right\rangle + \left\langle \gamma_1(t),\, \frac{d\gamma_2(t)}{dt} \right\rangle$$

b.

$$\frac{d}{dt}\left[\gamma_1(t), \gamma_2(t)\right] = \left[\frac{d\gamma_1}{dt}(t),\, \gamma_2(t)\right] + \left[\gamma_1(t),\, \frac{d\gamma_2}{dt}(t)\right]$$

c.

$$\frac{d}{dt}(\gamma_1(t), \gamma_2(t), \gamma_3(t)) = \left(\frac{d\gamma_1}{dt}(t), \gamma_2(t), \gamma_3(t)\right) + \left(\gamma_1(t), \frac{d\gamma_2}{dt}(t), \gamma_3(t)\right)$$

$$+ \left(\gamma_1(t), \gamma_2(t), \frac{d\gamma_3}{dt}(t)\right) = (\dot{\gamma}_1, \gamma_2, \gamma_3) + (\gamma_1, \dot{\gamma}_2, \gamma_3) + (\gamma_1, \gamma_2, \dot{\gamma}_3)$$

donde (, ,) es el triple producto escalar.

◁ Son igualmente cálculos directos y se omiten. ▷

EJEMPLO. Sean las curvas espaciales suaves

$$\gamma_1(t) = (\cos t, \operatorname{sen} t, 2\pi(t)) \quad \text{y} \quad \gamma_2(t) = (4t^2, 2t, -t)$$

$\triangleleft$ Ya que $\dot{\gamma}_1(t) = (-\text{sen}t, \cos t, \pi)$ y $\dot{\gamma}_2(t) = (8t, 2, -1)$ se tiene que

a.

$$\frac{d}{dt} < \gamma_1, \gamma_2 > = < \dot{\gamma}_1, \gamma_2 > + < \gamma_1, \dot{\gamma}_2 >$$

$$= < (-\text{sen}\, t, \cos t, 2\pi), (4t^2, 2t, -t) > + < (\cos t, \text{sen}\, t, 2\pi t), (8t, 2, -1) >$$

$$= -4t^2 \text{sen}\, t + 2t \cos t - 2\pi t + 8t \cos t + 2\text{sen}t - 2\pi t$$

$$= (2 - 4t^2)\text{sen}\, t + 10t \cos t - 4\pi t$$

b. De igual forma,

$$\frac{d}{dt}[\gamma_1, \gamma_2] = [\dot{\gamma}_1, \gamma_2] + [\gamma_1, \dot{\gamma}_2]$$

$$= \begin{vmatrix} e_1 & e_2 & e_3 \\ -\text{sen}\, t & \cos t & 2\pi \\ 4t^2 & 2t & -t \end{vmatrix} + \begin{vmatrix} e_1 & e_2 & e_3 \\ \cos t & \text{sen}\, t & 2\pi t \\ 8t & 2 & -1 \end{vmatrix}$$

$$= (-t \cos t - 4\pi t, -t\text{sen}\, t - 8\pi t^2, -2t\text{sen}\, t - 4t^2 \cos t)$$

$$+(-\text{sen}\, t - 4\pi, \cos t + 8\pi t^2, 2\cos t - 8t\text{sen}\, t)$$

$$= (-t \cos t - \text{sen}\, t - 8\pi t, \cos t - t\text{sen}\, t, -4t^2 \cos t - 10t\text{sen}\, t + 2\cos t) \quad \triangleright$$

Introducimos ahora el concepto de curva regular parametrizada, que es de gran utilidad en las aplicaciones a la mecánica, física e ingeniería. Este concepto, aunque en apariencia tiene definición abstracta, en realidad es establecida acorde a la experiencia cotidiana. [2]

DEFINICIÓN. Una **curva regular parametrizada** Γ es un subconjunto de $\mathbb{R}^3$ (ó de $\mathbb{R}^2$) tal que para cualquier punto $p \in \Gamma$ existe una función vectorial de variable real suave

$$\gamma : J \to \mathbb{R}^3$$

satisfaciendo que, si $J = (-\epsilon, \epsilon)$ para un número $\epsilon > 0$, entonces

a. La imagen de 0 bajo γ es el punto p, es decir,

$$\gamma(0) = p$$

b. La velocidad de γ no se anula en todo el intervalo, esto es,

$$\dot{\gamma}(t) \neq 0 \quad \text{para todo} \quad t \in J$$

[2]Por ejemplo, cuando se estudia una trayectoria espacial de un cuerpo, es difícil estudiarla en su totalidad, así que se procede a estudiarla por fragmento, es decir, por partes que se pueden parametrizar con una variable temporal adecuada, en un intervalo temporal adecuado.

c. La imagen de J bajo γ está contenida en Γ, en otras palabras,

$$\gamma(t) \in \Gamma \quad \text{para todo} \quad t \in J$$

Si la imagen de J bajo γ cubre a Γ, diremos que la pareja (γ, J) es una **parametrización** de Γ. En caso contrario, diremos que la pareja (γ, J) es sólo una parametrización de Γ alrededor del punto $p \in \Gamma$.

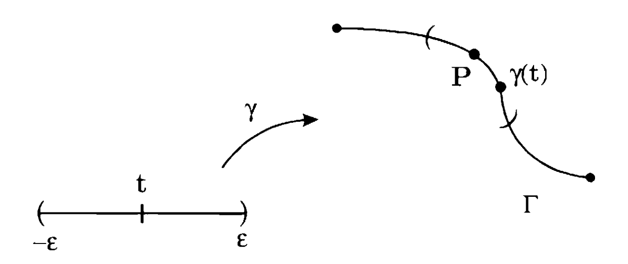


Figura 3.8: Curva regular parametrizada.

En coordenadas (x, y, z) del espacio cartesiano, la curva $\gamma(t)$ está definida por el vector

$$\gamma(t) = (x(t),\, y(t),\, z(t)) \quad t \in J,$$

donde las coordenadas $x(t)$, $y(t)$, $z(t)$ son funciones reales suaves definidas en el intervalo $J = (-\epsilon, \epsilon)$ que contiene a $t = 0$ (véase la figura 3.8).

En adelante, cada vez que hagamos referencia al dominio de parametrización de una curva Γ alrededor de un punto p lo definiremos por $[a, b]$ en lugar de $(-\epsilon, \epsilon)$, tomando en este caso, el dominio más amplio para parametrizar tal curva Γ.

Esto es, cada vez que demos una curva Γ, la entenderemos, sin pérdida de generalidad, como la imagen de una función

$$\gamma : [a, b] \to \Gamma$$

donde $[a, b] \subset \mathbb{R}$, siendo γ una función suave con coordenadas suaves.

Para $\gamma : [a, b] \to \Gamma$ curva regular parametrizada, en cada punto $\gamma(t) \in \Gamma$ se define un **vector tangente** a Γ en tal punto, como el vector derivada de γ en t, esto es,

$$\dot{\gamma} = v(t) = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right)$$

La notación $\dot{\gamma} = \dot{\gamma}(t)$ es muy conveniente, como ya hemos mencionado, a la hora de hacer cálculos, sobre todo en problemas de aplicación.

A continuación damos ejemplos de parametrizaciones de curvas $\Gamma \subset \mathbb{R}^3$ (o $\mathbb{R}^2$).

EJEMPLO. Considere la elipse en el plano con semieje mayor $a > 0$ y semieje menor b, definida por la ecuación

$$\Gamma : \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

$\lhd$ Una deformación del círculo unitario

$$x^2 + y^2 = 1,$$

parametrizado por $t \longrightarrow (\cos t,\, \operatorname{sen} t)$, por un escalar a en la primera coordenada, y por el escalar b en la segunda coordenada, nos da la parametrización requerida.

Esto es,

$$\Gamma : \begin{cases} x = a\,\cos t \\ y = b \operatorname{sen} t, \end{cases} \quad t \in \mathbb{R}$$

parametriza suavemente a la elipse Γ. Si escribimos $\gamma(t) = (a\cos t,\, b\operatorname{sen} t)$, entonces el vector tangente a γ en el punto $\gamma(t) \in \Gamma$ se calcula por

$$v(t) = \dot{\gamma}(t) = (-a\operatorname{sen} t,\, b\cos t)$$

Como $v(t) \neq 0$, entonces Γ resulta ser una curva regular. $\rhd$

EJEMPLO. Sea $\Gamma \subset \mathbb{R}^3$ una hélice contenida en un cilindro de radio 4 y con paso de altura $h = 4\pi$.

$\lhd$ Para poder parametrizar a Γ, partimos de que una hélice contenida en un cilindro de radio R y paso de altura $h = 2k\pi$ tiene la parametrización general

$$\dot{\gamma} = v(t) = t \longrightarrow (R\cos t,\, R\operatorname{sen} t, kt), \quad t \in \mathbb{R}$$

Si $R = 4$ y $h = 4\pi = 2(2\pi)$, entonces $k = 2$, lo que lleva a que la parametrización de Γ está bien definida por

$$\Gamma : \begin{cases} x = 4\cos t \\ y = 4\operatorname{sen} t \\ z = 2t \end{cases}$$

Si se define por $\gamma(t) = (4\cos t,\, 4\operatorname{sen} t,\, 2t)$, $t \in \mathbb{R}$, a tal curva, entonces el vector tangente en el punto $\gamma(t) \in \Gamma$ está descrito por

$$v(t) = \dot{\gamma}(t) = (-4\operatorname{sen} t,\, 4\cos t, 2),$$

vector que no se anula. Esto hace que la hélice Γ sea una curva regular. $\rhd$



EJEMPLO. Sea Γ el círculo de radio $R > 0$ con centro en el origen de coordenadas dado por

$$\Gamma : \quad \{(x,y) |\, x^2 + y^2 = R^2\}$$

$\lhd$ Al considerar la función vectorial $\gamma : (-\infty, \infty) \to \mathbb{R}^2$ dada por

$$\gamma(t) = (R\cos t, R\,\mathrm{sen}\, t)$$

entonces su imagen Γ es el círculo dado. Su derivada

$$\dot\gamma(t) = (-R\mathrm{sen}t, R\cos t)$$

es un vector que no se anula y, por lo tanto, se entenderá también como una **curva regular**.

Anotamos que la función vectorial $\tilde\gamma(t) = (R\cos 2t,\ R\sin 2t)$ tiene como imagen también al mismo círculo, pero su vector derivada es más grande de norma que la de $\gamma(t)$. De hecho,

$$\dot{\tilde\gamma}(t) = (-2R\,\mathrm{sen}\, 2t, 2R\cos 2t).$$

lo que implica que

$$||\dot{\tilde\gamma}(t)|| = ||(-2R\,\mathrm{sen}\, 2t,\ 2R\cos 2t)|| = 2R$$

$$= 2||(-R\,\mathrm{sen}\, t,\ R\cos t)|| = 2||\dot\gamma(t)||$$

lo que nos dice que la norma de $\dot{\tilde\gamma}(t)$ es el doble de la norma de $\dot\gamma(t)$. (La figura 3.9 ilustra esta situación). Decimos simplemente que la parametrización $\tilde\gamma(t)$ "va más rápida" que la de $\gamma(t)$. $\rhd$

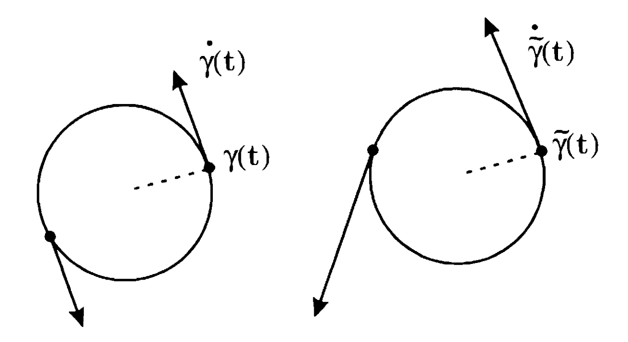


Figura 3.9: Parametrizaciones del círculo de radio $R > 0$.



3.2 La segunda derivada. Aceleración

Ya se ha mencionado sobre la clase de diferenciabilidad de una función vectorial de variable real (curva suave)

$$\gamma : J \to \mathbb{R}^3$$

relegando las definiciones a las propiedades de las funciones coordenadas

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases} , \quad t \in J$$

Particularmente, una curva de clase C^2 es una función vectorial de variable real cuyas componentes son funciones de clase C^2.

Entendemos por $\ddot{\gamma}(t)$ la segunda derivada de la curva γ en el punto $\gamma(t)$, y en coordenadas le escribimos

$$\ddot{\gamma}(t) = (\ddot{x}(t), \ddot{y}(t), \ddot{z}(t)) = \ddot{x}(t)e_1 + \ddot{y}(t)e_2 + \ddot{z}(t)e_3$$

donde los puntos indican derivación respecto a la variable temporal $t \in J$.

DEFINICIÓN. Al vector espacial $\ddot{\gamma}(t)$ se le llamará el vector **aceleración** de la curva γ en el punto $\gamma(t)$.

EJEMPLO. Sea el círculo de radio R, con centro en el origen, parametrizado por,

$$\gamma(t) = (R\cos t,\ R\,\mathrm{sen} t),\ t \in \mathbb{R}$$

◁ Un cálculo simple nos muestra que

$$\dot{\gamma}(t) = (-R\mathrm{sen}\, t,\ R\cos t)$$

$$\ddot{\gamma}(t) = (-R\cos t, -R\,\mathrm{sen}\, t) = -R(\cos t,\ \mathrm{sen}\, t)$$

lo que nos dice que $\gamma(t) = -\ddot{\gamma}(t)$. Esto es, el vector de posición $\gamma(t)$ y su aceleración $\ddot{\gamma}(t)$ son paralelos con sentidos opuestos. ▷

EJEMPLO. Dada la curva suave

$$\gamma(t) = (e^{2t}, \mathrm{sen} 4t, \cos t)$$

calcular $\dot{\gamma}(t)$ y $\ddot{\gamma}(t)$.

◁ Directamente, obtenemos los vectores velocidad y aceleración,

$$\dot{\gamma}(t) = (2e^{2t}, 4\cos 4t, -\mathrm{sen}\, t)$$



$$\ddot{\gamma}(t) = (4e^{2t}, -16\,\text{sen}\,4t, -\cos t) \quad \triangleright$$

EJEMPLO. Dada la curva suave

$$\gamma(t) = (t^2 + 1, t^3 - 2, t^4 + 1)$$

calcular $< \dot{\gamma}(t), \ddot{\gamma}(t) >$ y $[\dot{\gamma}(t), \ddot{\gamma}(t)]$.

$\triangleleft$ Al calcular directamente los vectores velocidad y aceleración de γ en $\gamma(t)$, obtenemos

$$\dot{\gamma}(t) = (2t, 3t^2, 4t^3)$$

$$\ddot{\gamma}(t) = (2, 6t, 12t^2)$$

de donde

$$< \dot{\gamma}(t), \ddot{\gamma}(t) >= 4t + 18t^3 + 48t^5$$

$$[\dot{\gamma}(t), \ddot{\gamma}(t)] = \begin{vmatrix} e_1 & e_3 & e_3 \\ 2t & 3t^2 & 4t^3 \\ 2 & 6t & 12t^2 \end{vmatrix} = e_1(36t^4 - 24t^4) - e_2(24t^3 - 8t^3)$$

$$+e_3(12t^2 - 6t^2) = (12t^4, -16t^3, 6t^2) \quad \triangleright$$

Consideremos ahora cualquier curva $\gamma = \gamma(t)$ tal que tenga una norma constante (por ejemplo $\gamma(t) = (R\cos t, R\text{sen}\,t)$ el círculo de radio R), esto es, para todo $t \in J$, tenemos que

$$||\gamma(t)|| = k$$

donde k es una constante no negativa.

$\triangleleft$ Ya que $k = ||\gamma(t)||$, entonces

$$k^2 = ||\gamma(t)||^2 =< \gamma(t), \gamma(t) >$$

de donde, al derivar cada miembro respecto al tiempo t obtenemos

$$0 = \frac{d}{dt} < \gamma(t), \gamma(t) >=< \dot{\gamma}, \gamma(t) > + < \gamma(t), \dot{\gamma}(t) >$$

$$= 2 < \dot{\gamma}(t), \gamma(t) >$$

es decir,

$$< \dot{\gamma}(t), \gamma(t) >= 0$$

lo que nos dice que $\gamma(t)$ y $\dot{\gamma}$ son ortogonales.

Así, utilizando este razonamiento para aquellas curvas con vectores velocidad con norma constante $||\dot{\gamma}(t)|| = k$ se tiene,



Las curvas suaves que tienen rapidez constante son tales que sus vectores velocidad y aceleración son ortogonales.

EJEMPLO. Sea la curva circular espacial

$$\gamma(t) = (4\cos t, 4\,\text{sen}\, t, 6)$$

◁ El vector velocidad se calcula

$$\dot{\gamma}(t) = (-4\,\text{sen}\, t, 4\cos t, 0)$$

cuya norma es $||\dot{\gamma}(t)|| = 4$.

Por lo tanto, el vector velocidad $\dot{\gamma}(t)$ es ortogonal al vector aceleración

$$\ddot{\gamma}(t) = (-4\cos r, -4\text{sen} t, 0)$$

debido a la discusión anterior, y como se puede apreciar mediante el cálculo directo,

$$< \dot{\gamma}(t), \ddot{\gamma}(t) >= 16\cos t\,\text{sen} t - 16\cos t\,\text{sen}\, t = 0 \quad \triangleright$$

Ejercicios

1. Dibujar la gráfica en $\mathbb{R}^2$ o $\mathbb{R}^3$ según corresponda, de las funciones

a. $\gamma_1(t) = (3t, 4\text{sen} t)$

b. $\gamma_2(t) = (\cos t, t + 2\text{sen} t)$

c. $\gamma_3(t) = (t^2 - 2t, t^2 + 2t)$

d. $\gamma_4(t) = (e^t \cos t, e^t \text{sen} t)$

e. $\gamma_5(t) = \left(\frac{1}{t}\cos t, \frac{1}{t}\text{sen}\, t\right)$

f. $\gamma_6(t) = \left(\frac{1+t}{t}\cos t, \frac{1+t}{t}\text{sen}\, t\right), \quad t \geq 1$

g. $\gamma_7(t) = (\cos t, \text{sen}\, t, t)$

h. $\gamma_8(t) = (4\cos t, 4\text{sen}\, t, e^t)$

2. a. Calcule la derivada de las funciones del ejercicio 1.

b. Calcular el vector tangente a cada curva en los puntos $t = 1, t = \pi/2, t = \pi$.

3. Calcular, para las curvas del ejercicio 1,

a. $\frac{d}{dt} < \gamma_1(t), \gamma_2(t) >$

b. $\frac{d}{dt}[\gamma_3(t), \gamma_4(t)]$

c. $\frac{d}{dt}(\gamma_1(t), \gamma_5(t), \gamma_6(t))$.

4. Calcular $\ddot{\gamma}(t)$ para cada curva en **1.** y el vector aceleración en los puntos $t = 1$, $t = \frac{\pi}{2}$ y $t = \pi$.

Capítulo 4

Campos escalares en $\mathbb{R}^3$

En este capítulo hacemos el análisis cualitativo de una función de varias variables reales, generalizando algunos conceptos conocidos ya para funciones reales de una variable real.

4.1 Regiones en $\mathbb{R}^2$ y $\mathbb{R}^3$

Iniciamos esta sección describiendo los objetos geométricos en $\mathbb{R}^n$ que van a ser utilizados a lo largo del capítulo. Éstos objetos gozan de características que son usadas para estudiar las propiedades de funciones que están definidas en subconjuntos del espacio euclidiano $\mathbb{R}^n$.

En general, en la teoría diferenciable de funciones reales de variable real, los dominios se presumen de la forma

$$D = (a, b)$$

es decir, intervalos **abiertos**.

Al intervalo (a, b) se le llama **abierto** pues no contiene las orillas a, b. Generalizamos este concepto de **conjunto abierto** en $\mathbb{R}^3$.

Si $p \in \mathbb{R}^n$ y $\epsilon > 0$ es un número real positivo, se define la **bola** de radio ϵ y centro en p, como el conjunto sólido

$$B_\epsilon(p) = \{q \in \mathbb{R}^n \mid ||q - p|| < \epsilon\}$$

DEFINICIÓN. Dado un conjunto $\Omega \subset \mathbb{R}^n$, se dice que es un **conjunto abierto** en $\mathbb{R}^n$, si para cada $p \in \Omega$ existe un número $\epsilon > 0$ tal que la bola de radio ϵ y centro en p está contenida totalmente en Ω, esto es,

$$B_\epsilon(p) \subset \Omega$$

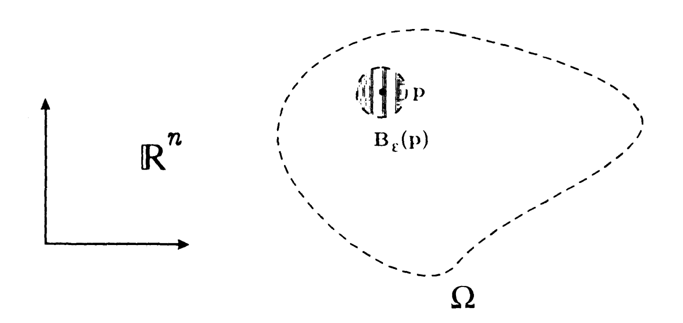


Figura 4.1: Conjunto abierto en $\mathbb{R}^n$.

Véase la figura 4.1.

EJEMPLO. Todo intervalo abierto $(a, b) \subset \mathbb{R}$ es un intervalo abierto.

◁ Esto ya se demostró anteriormente, notando que en el caso unidimensional una bola corresponde a un intervalo.

$$B_\epsilon(p) = \{q \in \mathbb{R} |\, ||q - p|| \, < \epsilon\} = \{q \in \mathbb{R} |\, |q - p| \, < \epsilon\}$$

$$= \{q \,|\, -\epsilon < q - p < \epsilon\} = \{q | \, p - \epsilon < q < p + \epsilon\}$$

$$= (p - \epsilon, p + \epsilon) \quad \triangleright$$

Los siguientes ejemplos omiten las demostraciones de las afirmaciones mostrando apenas una ilustración de los subconjuntos que intervienen. En general hacemos hincapié que un conjunto abierto Ω de $\mathbb{R}^n$ es un conjunto "gordito" que no tiene orilla (frontera).

EJEMPLO. Una bola de radio ϵ con centro en cualquier $p \in \mathbb{R}^n$ es un conjunto abierto.

◁ En este caso Ω es una bola $B_\epsilon(p)$ ▷

EJEMPLO. Sea $\Omega \subset \mathbb{R}^2$ la parte del plano definido por

$$\Omega = \{(x, y) | \, x > y\}$$

◁ El conjunto Ω es abierto en $\mathbb{R}^2$ (véase la figura 4.2). ▷

EJEMPLO. Sea $\Omega \subset \mathbb{R}^2$ el subconjunto dado por

$$\Omega = \{(x, y) | \, x^2 + y^2 > 4\}$$



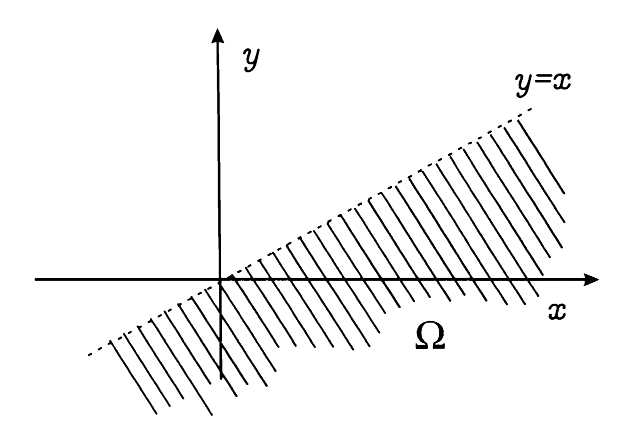


Figura 4.2: Semiplano inferior $x > y$.

$\lhd$ Claramente Ω es un subconjunto abierto en $\mathbb{R}^2$. $\rhd$

EJEMPLO. Sea el conjunto

$$\Omega = \{(x, y, z) | \, x^2 + y^2 < z^2\} \subset \mathbb{R}^3$$

que define el interior de un cono. Se afirma que es un conjunto abierto en $\mathbb{R}^3$.

$\lhd$ Al considerar la igualdad

$$x^2 + y^2 = z^2$$

tal ecuación describe a un cono y Ω define el interior del cono como lo muestra la figura 4.3. $\rhd$

EJEMPLO. Sea el conjunto $\Omega = \{(x, y, z) \,|\, x^2 + y^2 > 1\} \subset \mathbb{R}^3$.

$\lhd$ Entonces Ω es un conjunto abierto en $\mathbb{R}^3$ y es el exterior del cilindro unitario (véase la figura 4.4). $\rhd$

DEFINICIÓN. Definimos al conjunto $C \subset \mathbb{R}^n$ como un **conjunto cerrado** en $\mathbb{R}^n$, si su complemento en $\mathbb{R}^n$ es abierto.

Para ejemplos de conjuntos cerrados tenemos a los complementos de los conjuntos abiertos dados en los ejemplos anteriores.

EJEMPLO.

a. El conjunto

$$C = (-\infty, a] \cup [b, \infty)$$



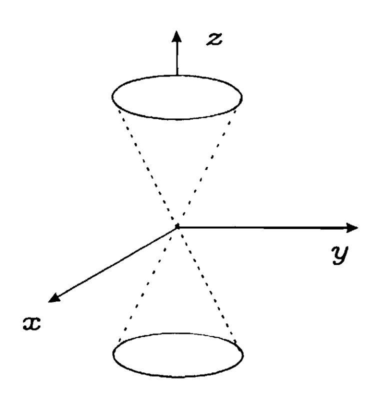


Figura 4.3: Interior del cono $z^2 = x^2 + y^2$.

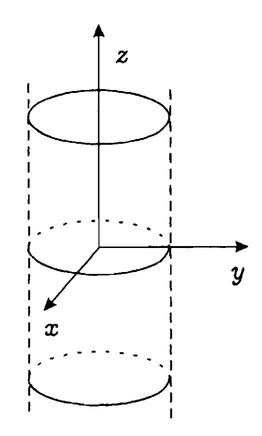


Figura 4.4: Exterior del cilindro unitario.

es cerrado en la recta real $\mathbb{R}$.

b. El conjunto

$$C = \{(x, y) \,|\, x \leq y\}$$

es cerrado en el plano $\mathbb{R}^2$.

c. El conjunto

$$C = \{(x, y) \,|\, x^2 + y^2 \leq 4\}$$

es cerrado en $\mathbb{R}^2$.

d. El exterior del cono, incluyéndolo

$$C = \{(x, y, z) \,|\, x^2 + y^2 \leq z^2\}$$

es un conjunto cerrado en $\mathbb{R}^3$.

e. El interior del cilindro unitario, incluyendo al cilindro

$$C = \{(x, y, z) \mid x^2 + y^2 \leq 1\}$$

es un conjunto cerrado de $\mathbb{R}^3$.

Es claro que existen subconjuntos de $\mathbb{R}^n$ que no son ni abiertos ni cerrados. Por ejemplo $D = (-1, 2]$ es un intervalo que no es abierto ni cerrado en $\mathbb{R}$.

Sea $\Omega \subset \mathbb{R}^3$ un subconjunto. Se dice que el punto $q \in \mathbb{R}^3$ es **punto frontera** de Ω si toda bola con centro en p intersecta a Ω y al complemento de Ω.

Al conjunto contenido en $\mathbb{R}^n$ definido por

$$\partial\Omega = \{p \in \mathbb{R}^n | \, p \quad \text{es frontera de} \quad \Omega\}$$

se llamará **frontera** de Ω en $\mathbb{R}^n$.

Damos algunos ejemplos de fronteras de conjuntos.

EJEMPLO.

a. La frontera de $\Omega = (a, b)$ es el conjunto formado por la pareja de puntos,

$$\partial(a, b) = \{a, b\}$$

b. La frontera del conjunto cerrado $C = \{(x, y) | \, x \leq y\}$ es la misma que la de su complemento, el abierto $\Omega = \{(x, y) | \, x > y\}$ y es la recta

$$\partial C = \partial\Omega = \{(x, y) | \, x = y\}$$

Tal conjunto está contenido en el cerrado C y no pertenece al subconjunto abierto Ω.

c. La frontera del abierto $\Omega = \{(x, y) | \, x^2 + y^2 < 4\}$ es la misma que la del cerrado $C = \{(x, y) | \, x^2 + y^2 \leq 4\}$ y la constituye el conjunto cilíndrico

$$\partial\Omega = \partial C = \{(x, y) \mid x^2 + y^2 = 4\}$$

que está contenido en el cerrado C.

d. Dados los conjuntos $\Omega = (x, y, z) \mid x^2 + y^2 < z^2\}$ abierto, y $C = \{(x, y, z) \mid x^2 + y^2 \geq z^2\}$ cerrado, tenemos

$$\partial C = \partial\Omega = \{(x, y, z) \mid x^2 + y^2 = z^2\}$$

es la frontera tanto de C como de Ω.

Entre las características geométricas de un subconjunto de $\mathbb{R}^n$ está aquella que nos asegura que tal subconjunto no está dividido en pedazos ajenos entre sí.

DEFINICIÓN. Decimos que el conjunto $\Omega \subset \mathbb{R}^n$ es **conexo** si dados dos puntos $p, q, \in \Omega$, existe una curva Γ que los une, totalmente contenida en Ω.

Si Ω no es un conjunto conexo se llamará **disconexo**, (véase la figura 4.5).

EJEMPLO. Sea el conjunto del plano

$$\Omega = \{(x, y) | \, x^2 + y^2 \leq 1 \quad \text{ó} \quad x^2 + y^2 \geq 4\}$$

Se afirma que Ω es disconexo.

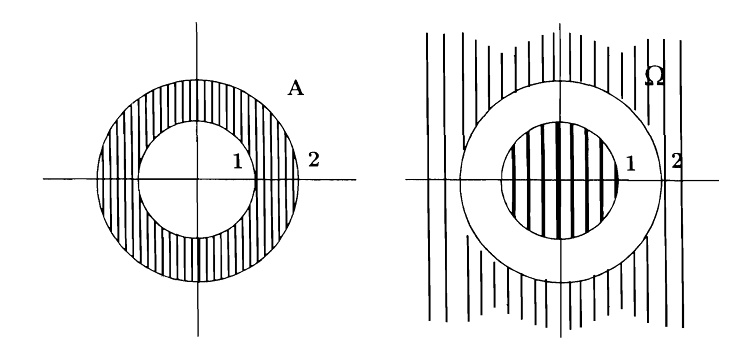


Figura 4.5: Conjuntos conexo y disconexo.

$\triangleleft$ Dados p y q con $||p|| \leq 1$ con $||q|| \geq 2$ se tiene cualquier curva que los une se sale del conjunto Ω. $\triangleright$

EJEMPLO. El conjunto anular

$$A = \{(x, y) \in \mathbb{R}^2 | \quad 1 \leq x^2 + y^2 \leq 4\}$$

encerrado entre los círculos de radios 1 y 2 respectivamente, con centros en el origen, es un conjunto conexo.

En general un conjunto conexo es aquél que está formado por una sola pieza. Esto es, un conjunto conexo en $\mathbb{R}^n$ se entiende intuitivamente como un conjunto que no se divida en varias piezas.

DEFINICIÓN. Por una **región** en $\mathbb{R}^n$ entenderemos un conjunto $\Omega \subset \mathbb{R}^n$ que es abierto y conexo.

4.2 Campos escalares en $\mathbb{R}^3$

Pasamos ahora a definir a una función real con argumento vectorial, utilizando los conceptos dados de abierto y conexo.

Por un **campo escalar** definido en la región $\Omega \subset \mathbb{R}^n$

$$f : \Omega \to \mathbb{R}$$

entendemos a la función real de variable vectorial, es decir, una asociación tal que cada punto $p \in \Omega$ se le asocia el número real $f(p)$.

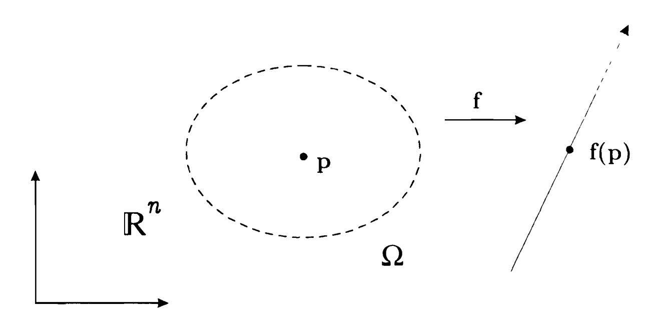


Figura 4.6: Función vectorial $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$.

En coordenadas x, y, z del espacio $\mathbb{R}^3$ esta asociación se escribe, para el punto $p = (x, y, z)$,

$$f(p) = f(x, y, z)$$

Esto es, si $(x, y, z) \in \Omega$, bajo la función f se le asociaría el número real

$$(x, y, z) \to f(x, y, z)$$

En adelante, usaremos indistintamente los nombres de función real de variable vectorial y campo escalar.

DEFINICIÓN. La **gráfica** de una función $f : \Omega \to \mathbb{R}$ es un subconjunto Graf(f) de $\mathbb{R}^4$ definido por

$$\text{Graf}\,(f) = \{(p,\ f(p))\,|\, p \in \Omega,\ f(p) \in \mathbb{R}\}$$

$$= \{((x, y, z),\ f(x, y, z))|\,(x, y, z) \in \Omega,\ f(p) \in \mathbb{R}\}$$

Análogamente, para una función definida en la región plana $\Omega \subset \mathbb{R}^2$, $f : \Omega \to \mathbb{R}$ se tiene su gráfica como el subconjunto

$$\text{Graf}\,(f) = \{(x, y, z)\,|\,(x, y) \in \Omega,\ z = f(x, y)\}$$

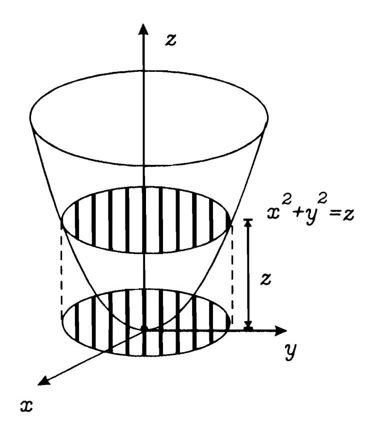


Figura 4.7: Gráfica de $f(x, y) = x^2 + y^2$

contenido en $\mathbb{R}^3$, constituyendo una superficie en tal espacio tridimensional.

Damos algunos ejemplos donde la regla de correspondencia es dada de antemano.

EJEMPLO. Sea el campo escalar cuya regla está dada por

$$f(x, y) = x^2 + y^2$$

$\triangleleft$ Claramente, el dominio de f es $\Omega = \mathbb{R}^2$ y su gráfica es la superficie definida por el conjunto contenido en $\mathbb{R}^3$

$$\begin{aligned} Graf\,(f) &= \{((x, y),\, x^2 + y^2) |\, (x, y) \in \mathbb{R}^2\} \\ &= \{(x, y, z) |\, z = x^2 + y^2\} \end{aligned}$$

que resulta ser un paraboloide circular (véase la figura 4.7). $\triangleright$

EJEMPLO. Sea la función $f : \Omega \to \mathbb{R}$ dada por la regla de correspondencia

$$f(x, y) = \sqrt{4 - x^2 - y^2}$$

$\triangleleft$ Para calcular el dominio Ω imponemos la condición $4 - x^2 - y^2 \geq 0$ o equivalentemente,

$$4 \geq x^2 + y^2 \Longleftrightarrow 2 \geq \sqrt{x^2 + y^2}$$

Esto es, para $(x, y) \in \Omega$ se debe cumplir que $||(x, y)|| \leq 2$ lo que nos dice que el dominio Ω deberá ser un disco cerrado de radio 2 con centro en el origen de coordenadas.

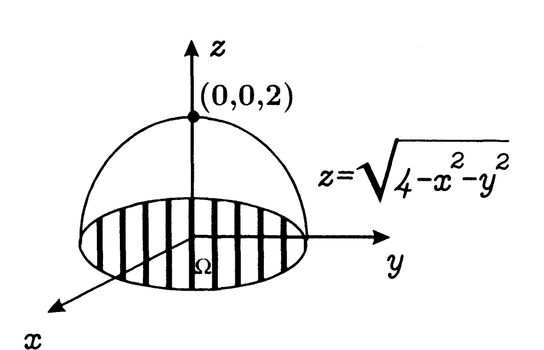


Figura 4.8: Gráfica de $f(x, y) = \sqrt{4 - x^2 - y^2}$.

La gráfica de f es el conjunto cerrado

$$\text{Graf}\,(f) = \{(x, y, \sqrt{4 - x^2 - y^2}) \mid x^2 + y^2 \leq 4\}$$

Si ponemos $z = \sqrt{4 - x^2 - y^2}$ entonces, al elevar cuadrados obtenemos

$$z^2 = 4 - x^2 - y^2$$

de donde,

$$x^2 + y^2 + z^2 = 4$$

que es una esfera de radio 2 con centro $(0, 0, 0)$.

De esta manera, la gráfica de f es la parte superior de tal esfera como lo muestra la figura 4.8. ▷

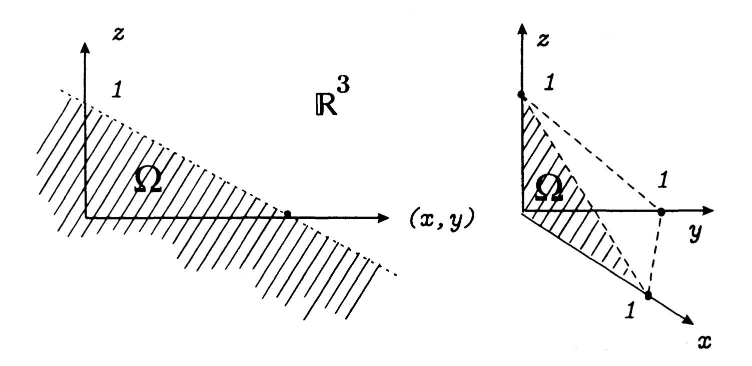


Figura 4.9: Parte del plano $x + y + z = 1$ en el primer octante.

EJEMPLO. Sea la función $f : \Omega \to \mathbb{R}$ dada por la regla

$$f(x, y, z) = \ln(1 - x - y - z)$$



$\triangleleft$ La función está definida en una región de $\mathbb{R}^3$ determinada por la relación

$$1 - x - y - z > 0$$

es decir,

$$1 > x + y + z$$

En la igualdad $x + y + z = 1$ tenemos un plano con cortes con los ejes $x = 1$, $y = 1$ y $z = 1$.

Al tomar el punto $p = (0,0,0)$ se tiene que sus coordenadas satisfacen

$$0 + 0 + 0 = 0 < 1$$

es decir, p satisface la relación $x + y + z < 1$ y está por abajo del plano mencionado. Es natural pensar que todos los puntos (x, y, z) satisfaciendo la relación $x + y + z < 1$ están situados por debajo de tal plano.

De esta manera el dominio de f es la región Ω de $\mathbb{R}^3$ formada por los puntos debajo del plano $x + y + z = 1$ (véase la figura 4.9).

4.3 Superficies y curvas de nivel

En esta parte estudiaremos un método para trazar aproximadamente la gráfica de una función de dos variables mediante los conjuntos de nivel de tal función (imágenes inversas) contenidos en el dominio.

DEFINICIÓN. Para un campo escalar $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ y un valor $c \in \mathbb{R}$ dados, se define el **conjunto de nivel** de f al valor c, definido por $f^{-1}(c)$, como la imagen inversa de c bajo f, esto es,

$$f^{-1}(c) = \{(x, y, z) | \, f(x, y, z) = c\}$$

En otras palabras, es el conjunto de puntos $(x, y, z) \in \Omega$ que bajo f alcanzan el mismo valor $f(x, y, z) = c$.

EJEMPLO. Sea la función

$$f(x, y, z) = x^2 + y^2 + z^2$$

Calcular los conjuntos de nivel para los valores $c = 1, 0, 1$.

$\triangleleft$ Para el caso $c = -1$, se tiene que por definición

$$f^{-1}(-1) = \{(x, y, z) | \, x^2 + y^2 + z^2 = -1\}$$

Esto es, los puntos (x, y, z) en tal conjunto deberán satisfacer la relación $x^2 + y^2 + z^2 = -1$, lo cual es imposible.

Consecuentemente, el conjunto de nivel obtenido $f^{-1}(-1) = \emptyset$ es el conjunto vacío.

Consideremos ahora el valor $c = 0$. Entonces en este caso,

$$f^{-1}(0) = \{(x, y, z) | \, x^2 + y^2 + z^2 = 0\}$$

es decir, los puntos contenidos en tal conjunto deberán satisfacer la relación $x^2 + y^2 + z^2 = 0$, y tal relación es apenas resuelta cuando $x = y = z = 0$.
Por lo tanto, el conjunto de nivel para $c = 0$ es apenas el punto

$$f^{-1}(0) = \{(0, 0, 0)\}$$

Finalmente, consideramos el valor $c = 1$, y entonces en este caso,

$$f^{-1}(1) = \{(x, y, z) | \, x^2 + y^2 + z^2 = 1\}$$

Es decir, los puntos en tal conjunto deberán satisfacer la relación $x^2 + y^2 + z^2 = 1$, la cual corresponde a la ecuación de la esfera $S^2 \subset \mathbb{R}^3$, esto es,

$$f^{-1}(1) = S^2$$

De esta última discusión, se observa que para un valor c arbitrario los puntos (x, y, z) en el conjunto de nivel $f^{-1}(c)$ deberán de satisfacer la ecuación

$$x^2 + y^2 + z^2 = c$$

la cual tiene sentido apenas (puede ser resuelta) para $c \geq 0$, es decir, $f^{-1}(c)$ es vacío si $c < 0$. Si $c = 0$ hemos ya notado que $f^{-1}(0)$ se reduce al punto $(0, 0, 0)$. Por otro lado, si $c > 0$ entonces $f^{-1}(c)$ es una esfera con centro en el origen y de radio $R = \sqrt{c}$ como lo muestra la figura 4.10.

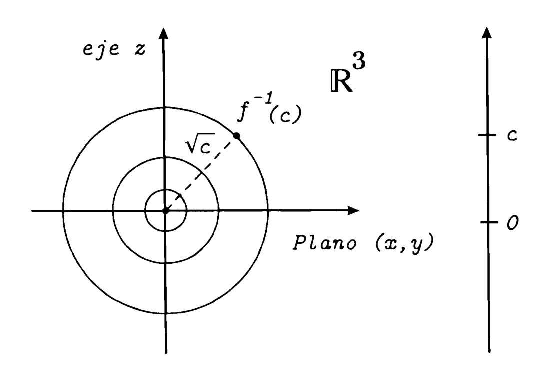


Figura 4.10: Conjuntos de nivel de $x^2 + y^2 + z^2 = c$

Notamos además que para $c_1 \neq c_2$ se tienen dos esferas diferentes $f^{-1}(c_1) \neq f^{-1}(c_2)$ en virtud de que los radios de cada una es diferente.

Por lo tanto, los conjuntos de nivel folían (dividen) al dominio de la función $f(x, y, z) = x^2 + y^2 + z^2$ mediante esferas (hojas) que no se intersectan entre sí. $\triangleright$

La afirmación última es válida en general para cualquier campo escalar $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$. Esto es,

LEMA 4.1 *Los conjuntos de nivel de un campo escalar f folían (dividen) al dominio mediante conjuntos (hojas) $f^{-1}(c)$ que no se intersectan entre sí.*

$\triangleleft$ Si para los valores $c_1 \neq c_2 \in \mathbb{R}$, los conjuntos de nivel respectivos $f^{-1}(c_1)$ y $f^{-1}(c_2)$ se intersectan, $f^{-1}(c_1) \cap f^{-1}(c_2) \neq \phi$, entonces sea $p \in f^{-1}(c_1) \cap f^{-1}(c_2)$. Ya que $p \in f^{-1}(c_1)$ y $p \in f^{-1}(c_2)$, entonces $f(p) = c_1$ y $f(p) = c_2$, lo que no puede ser pues f es una función y a cada punto del dominio le asocia un valor y sólo uno. $\triangleright$

La discusión que hemos realizado se puede llevar a cualquier función de varias variables, particularmente para funciones del tipo $f(x, y)$, donde los conjuntos de nivel $f^{-1}(c)$ son curvas planas que folían el dominio de f (véase la figura 4.11).

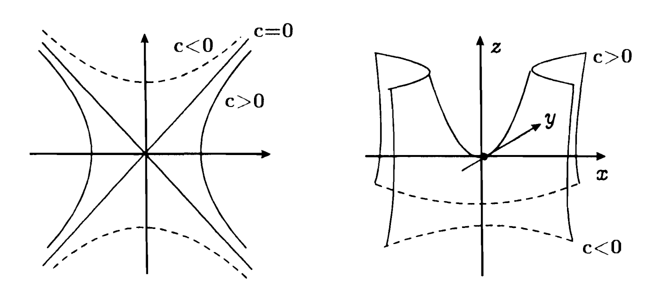


Figura 4.11: Curvas de nivel y gráfica de la función $f(x, y) = x^2 - y^2$.

EJEMPLO. Dada la función en dos variables,

$$f(x, y) = x^2 - y^2$$

describir a los conjuntos de nivel para $c = -1, 0, 1$.

$\triangleleft$ Para el valor $c = -1$, se tiene la relación

$$x^2 - y^2 = -1$$



que corresponde a una hipérbola con sus componentes verticales y pasando por los vértices $(0, 1)$ y $(0, -1)$ respectivamente.

Para el valor $c = 0$ se deberá de cumplir la relación

$$x^2 - y^2 = 0 \Longleftrightarrow (x, +y)(x - y) = 0 \Longleftrightarrow y = \pm x$$

que corresponde a un par de rectas por el origen de coordenadas.

Tales rectas son las asíntotas de la hipérbola encontrada cuando se estudio el nivel $c = -1$.

Finalmente para el caso $c = 1$, se cumplirá la relación

$$x^2 - y^2 = 1$$

que corresponde a una hipérbola con sus componentes horizontales que pasan por los vértices $(1, 0)$ y $(-1, 0)$ respectivamente, teniendo como rectas asíntotas a $y = \pm x$.

Un análisis semejante al realizado nos permite concluir que si $c \in \mathbb{R}$ es un valor arbitrario se cumple que

a. Para $c < 0$, el conjunto $f^{-1}(c)$ es una hipérbola vertical con vértices en los puntos $(0, \sqrt{c}), (0, -\sqrt{c})$, y con rectas asíntotas $y = \pm x$.

b. Para $c = 0, f^{-1}(c)$ es la pareja de rectas por el origen $y = \pm x$.

c. Para $c > 0$, el conjunto $f^{-1}(c)$ es una hipérbola horizontal con vértices en los puntos $(\sqrt{c}, 0), (-\sqrt{c}, 0)$, y con rectas asíntotas $y = \pm x$.

La figura 4.11 ilustra las curvas de nivel en este ejemplo. ▷

Observemos detenidamente el proceso de foliar el dominio Ω de la función f mediante conjuntos de nivel. Damos inicialmente un valor de f a la "altura" $c \in \mathbb{R}$ y buscamos el conjunto $f^{-1}(c)$ que es la imagen inversa de c bajo f. En otras palabras, $f^{-1}(c)$ es la forma del conjunto que tiene la gráfica de la función f cuando se observa sólo a la altura c de los posibles valores.

Más correctamente, cada subconjunto de nivel $f^{-1}(c)$ es la parte de la gráfica de la función f suspendida dentro de la gráfica a un nivel del valor igual a c, lo que nos dice que se puede construir la gráfica de la función $f : \Omega \to \mathbb{R}$ con el simple hecho de foliar Ω con los conjuntos de nivel, y después colocar cada conjunto de nivel a su altura correspondiente. Este proceso nos recuerda a la descripción de las zonas geográficas mediante mapas topográficos que detallan la configuración de ellas utilizando curvas de nivel. Cada curva de nivel corresponde a una altura determinada y el conjunto total de curvas nos ayuda a describir la topografía de la zona estudiada (véase figura 4.12).

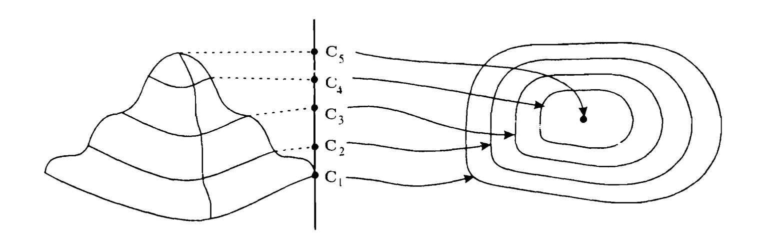


Figura 4.12: Curvas de nivel en topografía.

EJEMPLO. Consideremos nuevamente el campo escalar,

$$f(x, y) = x^2 - y^2$$

$\lhd$ Ya hemos estudiado cuáles son las curvas de nivel de esta función, por lo que se tienen los siguientes resultados.

a. Para un valor $c < 0$, el corte de la gráfica

$$\text{Graf}(f) = \{(x, y, z) | \, z = x^2 - y^2\}$$

con el plano horizontal (a la altura) $z = c$ corresponde a una hipérbola en el plano $\mathbb{R}^2_{x,y}$ en dirección del eje y, con vértices en $(0, \sqrt{c}), (0, -\sqrt{c})$ y asíntotas $y = \pm x$.

b. Para el valor $c = 0$, el corte de la gráfica Graf(f) con el plano horizontal $z = 0$ corresponde al par de rectas $y = \pm x$ en el plano $\mathbb{R}^2_{x,y}$.

c. Para el valor $c > 0$, el corte de la gráfica Graf(f) con el plano horizontal $z = 0$ corresponde a una hipérbola en el plano $\mathbb{R}^2_{x,y}$ en dirección del eje x, con vértices en $(\sqrt{c}, 0), (-\sqrt{c}, 0)$, y con asíntotas $y = \pm x$.

La figura 4.11 describe la gráfica de esta función que es un paraboloide hiperbólico (**silla de montar**). $\rhd$

Los siguientes ejemplos nos indican que en general, hacer el trazo de la gráfica de una función de dos variables no es simple cuando se usa la metodología de las curvas de nivel, y que hay que encontrar argumentos más sotisficados para el trazo, debido a que la estructura de las curvas de nivel es complicado, o demasiado simple para la elaboración de tal trazo.

EJEMPLO. Considérese la función definida en $\Omega = \mathbb{R}^2 - \{(0, 0)\}$ por la regla de correspondencia

$$f(x, y) = \frac{xy}{x^2 + y^2}$$



◁ Calculamos sus curvas de nivel para un valor arbitrario $c \in \mathbb{R}$. Los puntos $(x, y) \in f^{-1}(c)$ deberán satisfacer la relación

$$\frac{xy}{x^2 + y^2} = c$$

o equivalentemente (omitiendo el origen de coordenadas),

$$xy = c(x^2 + y^2) \Longleftrightarrow cx^2 - xy + cy^2 = 0$$

Si se resuelve la última ecuación para y en términos de x, se tiene que

$$y = \frac{x \pm \sqrt{x^2 - 4c^2x^2}}{2c} = \frac{1 \pm \sqrt{1 - 4c^2}}{2c}x$$

que corresponde a ecuaciones de rectas por el origen cuya pendiente depende del valor de c.

Tales valores de c están condicionados a que $1 - 4c^2 \geq 0$, es decir, a que $-\frac{1}{2} \leq c \leq \frac{1}{2}$. Las curvas de nivel, que realmente no pasan por $(0, 0)$, están trazadas en la figura 4.13 para ciertos valores de c. Claramente es difícil realizar el dibujo de la gráfica con estos elementos en virtud de que el haz de líneas obtenidas al ser levantados a sus valores respectivos dificultan su pegado continuo, además de la aparición del **punto singular** $p = (0, 0)$. La traza de la gráfica de f realizada por el programa de **Mathematica** se muestra en la figura 4.13 que describe su complejidad cerca del punto singular. ▷

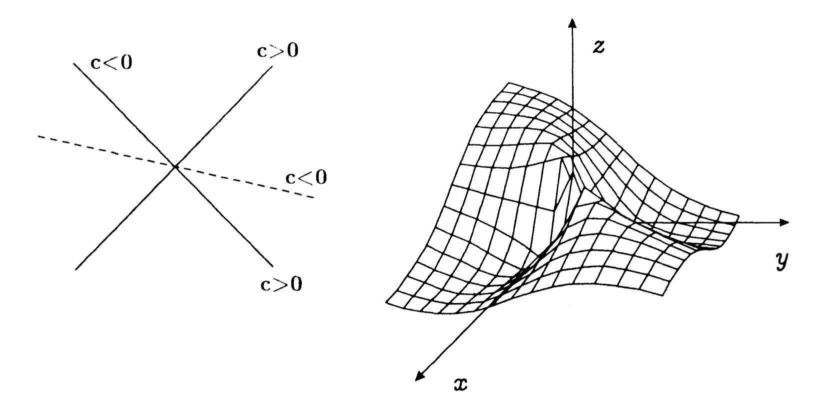


Figura 4.13: Curvas de nivel y gráficas de $f(x, y) = \frac{xy}{x^2+y^2}$.

EJEMPLO. Consideremos el campo escalar cuártico

$$f(x, y) = x^4 - y^4 - 4xy$$



$\triangleleft$ Para el valor $c \in \mathbb{R}$, los puntos $(x, y) \in f^{-1}(c)$ deberán satisfacer la ecuación

$$x^4 + y^4 - 4xy = c.$$

Tal ecuación no es a primera vista reconocida y, de hecho, la forma de las curvas de nivel varían según el valor c. En el capítulo 9 (de los ápendices) buscaremos tales curvas de nivel utilizando un análisis cualitativo de la función alrededor de los llamados **puntos críticos**, que determinan localmente (alrededor de ellos) la estructura de las curvas de nivel. De hecho mostraremos las siguientes propiedades de la función f.

a. Si $c < -2$ entonces $f^{-1}(c)$ es vacío, es decir, sólo cuando $c \geq -2$ el conjunto de nivel $f^{-1}(c)$ es no vacío.

b. Cuando $c = -2$ la curva de nivel es la pareja de **puntos críticos** aislados

$$f^{-1}(-2) = \{(-1, -1), (1, 1)\}$$

c. Cuando $-2 < c < 0$, la curva de nivel $f^{-1}(c)$ está compuesta por una pareja de curvas cerradas que encierran a los puntos críticos mencionados en b.

d. Cuando $c = 0$ la curva de nivel $f^{-1}(0)$ es una curva que se autointersecta en $(0, 0)$ en forma de "ocho" y que en cada uno de sus pétalos encierra las parejas de curvas mencionados en c.

e. Cuando $c > 0$ la curva de nivel $f^{-1}(c)$ es un circuito cerrado que encierra a la figura de "ocho" dada por $f^{-1}(0)$.

Cabe resaltar que una prueba contundente de todos los resultados mencionados en este ejemplo serían provistos por la teoría básica de la rama de las matemáticas conocida como **Geometría Algebraica**.

La figura 4.14 ilustra las curvas de nivel de la función y su gráfica trazada por el programa de **Mathematica**. $\triangleright$

La moraleja de los últimos ejemplos es que, en general, no es simple ni siquiera construir las curvas de nivel de una función .

Más aún, ya teniendo las curvas de nivel de una función (vértebras de la gráfica), en muchas ocasiones no es fácil construir su gráfica (columna vertebral).

Lo recomendable, entonces, durante el análisis de un campo escalar de dos variables, es el uso de un programa de graficación ejecutado por un ordenador.

Para el caso de una función de tres variables, la construcción de su gráfica es imposible debido a que está sumergida en un espacio real de dimensión cuatro, objeto que no es simple dibujar.



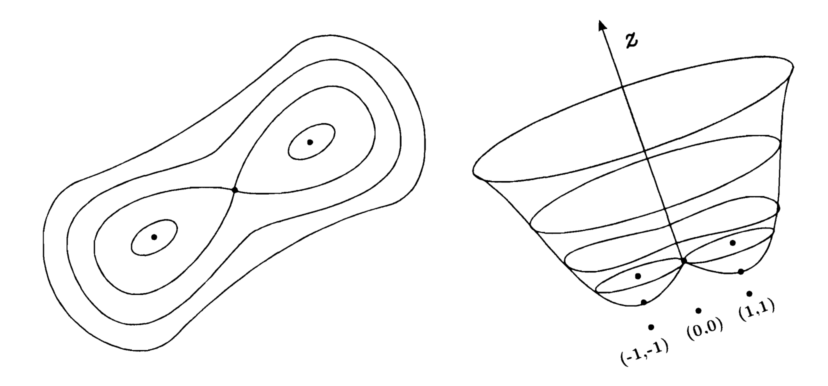


Figura 4.14: Curvas de nivel y gráfica de $f(x, y) = x^4 + y^4 - 4xy$.

4.4 Derivadas parciales y el gradiente

A continuación iniciamos el estudio de la propiedad de diferenciabilidad de un campo escalar definido en una región del espacio $\mathbb{R}^3$ (ó $\mathbb{R}^2$).

Sean, $f : \Omega \to \mathbb{R}$ una función real de variable vectorial definida en la región $\Omega \subset \mathbb{R}^3$, y $p = (x_0, y_0, z_0) \in \Omega$ un punto arbitrario. Si dejamos fijas las coordenadas y_0, z_0 y variamos la variable x_0 en el intervalo $[x_0, x_0 + h]$ podemos formar el cociente (**parcial**)

$$\frac{f(x_0 + h,\, y_0, z_0) - f(x_0, y_0, z_0)}{h}$$

y tomar el límite cuando $h \to 0$ (véase la figura 4.15).

DEFINICIÓN. Si tal límite existe, se llamará la **derivada parcial** de f respecto a x en el punto $p = (x_0, y_0, z_0)$ y se entenderá por

$$\frac{\partial f}{\partial x}(x_0, y_0, z_0) = \lim_{h \to 0} \frac{f(x_0 + h,\, y_0, z_0) - f(x_0, y_0, z_0)}{h}$$

Tal cantidad define la **variación** de la función f en el punto $p = (x_0, y_0, z_0)$ en la dirección del eje x.

Análogamente, se construyen la **derivada parcial** de f respecto a y en $p = (x_0, y_0, z_0)$ como

$$\frac{\partial f}{\partial y}(x_0, y_0, z_0) = \lim_{k \to 0} \frac{f(x_0, y_0 + k, z_0) - f(x_0, y_0, z_0)}{k}$$

y la **derivada parcial** de f respecto a z en el punto (x_0, y_0, z_0) mediante la igualdad

$$\frac{\partial f}{\partial z}(x_0, y_0, z_0) = \lim_{\ell \to 0} \frac{f(x_0, y_0, z_0 + \ell) - f(x_0, y_0, z_0)}{\ell}$$



si tales límites existen.

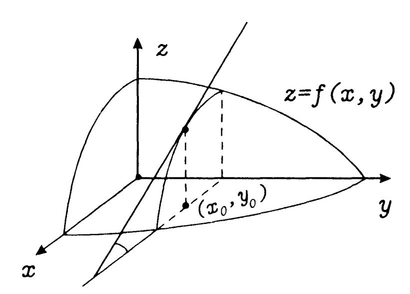


Figura 4.15: Derivada parcial de una función respecto a x.

Como (x_0, y_0, z_0) es arbitrario, el cálculo de tales parciales se obtiene derivando ordinariamente la función dada respecto a la variable indicada dejando fijas las otras variables.

EJEMPLO. Sea el campo escalar

$$f(x, y, z) = x^2 + y^2 + z^2$$

◁ Calculamos sus derivadas parciales derivando ordinariamente respecto a la variable indicada, fijando las otras variables, obteniendo en cualquier punto arbitrario (x, y, z) del plano,

$$\frac{\partial f}{\partial x}(x, y, z) = 2x, \quad \frac{\partial f}{\partial y}(x, y, z) = 2y, \quad \frac{\partial f}{\partial z}(x, y, z) = 2z \quad \triangleright$$

EJEMPLO. Sea la función en tres variables

$$f(x, y, z) = e^{xyz}$$

◁ Por un cálculo directo se tiene, en cualquier punto (x, y, z) del dominio,

$$\frac{\partial f}{\partial x}(x, y, z) = yz\, e^{xyz}$$

$$\frac{\partial f}{\partial y}(x, y, z) = xz\, e^{xyz}$$

$$\frac{\partial f}{\partial z}(x, y, z) = xy\, e^{xyz} \quad \triangleright$$

EJEMPLO. Dado el campo escalar

$$f(x, y) = \operatorname{sen}(xy)$$

calcular las derivadas parciales de f en el punto $p = (1, \pi)$.

$\lhd$ Calculamos directamente en cada punto (x, y) del dominio,

$$\frac{\partial f}{\partial x}(x, y) = y \cos(xy)$$

$$\frac{\partial f}{\partial y}(x, y) = x \cos(xy)$$

y al evaluar en $p = (1, \pi)$ se tiene,

$$\frac{\partial f}{\partial x}(1, \pi) = \pi \cos \pi = -\pi$$

$$\frac{\partial f}{\partial y}(1, \pi) = 1 \cos \pi = -1 \quad \rhd$$

En ocasiones, para fines de cálculo, es conveniente omitir el argumento donde se toman las derivadas parciales, esto es,

$$\frac{\partial f}{\partial z}(x, y, z) \quad \text{y} \quad \frac{\partial f}{\partial z}$$

se sobreentiende que identifican a la misma cantidad.

Para el campo escalar $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ se define en el punto $p \in \Omega$ el vector formado por las derivadas parciales en tal punto, llamado el **gradiente** del campo f en el punto p, y definido por

$$\nabla f_p = \left(\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p), \frac{\partial f}{\partial z}(p) \right)$$

EJEMPLO. Si $f(x, y) = \operatorname{sen}(xy)$ y $p = (1, \pi)$ entonces el vector gradiente de f en p es el vector,

$$\nabla f_{(1,\pi)} = \left(\frac{\partial f}{\partial x}(1, \pi), \frac{\partial f}{\partial y}(1, \pi) \right) = (-\pi, -1)$$

EJEMPLO. Sean $f(x, y) = x^2 y^3$ y $p = (3, 2)$.

$\lhd$ Para cualquier punto P se tiene

$$\nabla f_p = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (2xy^3, 3xy^2)_p$$

y, por lo tanto, el gradiente de f en el punto $p = (3, 2)$ es el vector

$$\nabla f_{(3,2)} = (2xy^3, 3xy^2)_{(3,2)} = (48, 108) \quad \rhd$$

Tenemos el siguiente resultado acerca de las propiedades aritméticas del gradiente de un campo escalar.

TEOREMA 4.1 *Sean $f, g : \Omega \subset \mathbb{R}^3 - \mathbb{R}$ dos campos escalares tales que sus derivadas parciales existan en cada punto, entonces*

a. *El gradiente de la suma de funciones satisface la regla,*

$$\nabla(f+g)_p = \nabla f_p + \nabla g_p$$

b. *El gradiente es lineal respecto a escalares, es decir,*

$$\nabla(\lambda f)_p = \lambda \nabla f_p$$

donde $\lambda \in \mathbb{R}$ es un escalar.

c. *Se cumple La* **fórmula de Leibniz**

$$\nabla(fg) = g\nabla f + f\nabla g$$

$\lhd$ Damos apenas la demostración de **a.**

$$\begin{aligned}
\nabla(f+g)_p &= \left(\frac{\partial(f+q)}{\partial x}(p), \frac{\partial(f+g)}{\partial y}(p), \frac{\partial(f+g)}{\partial z}(p)\right) \\
&= \left(\frac{\partial f}{\partial x}(p) + \frac{\partial g}{\partial x}(p), \frac{\partial f}{\partial y}(p) + \frac{\partial g}{\partial y}(p), \frac{\partial f}{\partial z}(p) + \frac{\partial g}{\partial z}(p)\right) \\
&= \left(\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p), \frac{\partial f}{\partial z}(p)\right) + \left(\frac{\partial g}{\partial x}(p), \frac{\partial g}{\partial y}(p), \frac{\partial g}{\partial z}(p)\right) \\
&= \nabla f_p + \nabla g_p \quad \rhd
\end{aligned}$$

EJEMPLO. Sea la función continua en tres variables dada por,

$$f(x,y,z) = \operatorname{sen} xy + e^{xyz}$$

$\lhd$ Entonces, para cada punto $p = (x, y, z)$ en el espacio,

$$\begin{aligned}
\nabla f_p &= \nabla(\operatorname{sen} xy)_p + \nabla(e^{xyz})_p \\
&= (y\cos xy, x\cos xy, 0)_p + (yze^{xyz}, xze^{xyz}, xye^{xyz})_p \\
&= (y\cos xy + yze^{xyz}, x\cos xy + xze^{xyz}, xye^{xyz})_p
\end{aligned}$$

Particularmente, en el punto $p = (1, \pi, 0)$ se tiene

$$\begin{aligned}
\nabla f_{(1,\pi,0)} &= (y\cos xy + yze^{xyz}, x\cos xy + xze^{xyz}, xye^{xyz})_{(1,\pi,0)} \\
&= (\pi\cos\pi) + \pi(0)e^{(1)(\pi)(0)}, 1\cos(\pi) + (1)(0)e^0, \pi e^0) \\
&= (-\pi, -1, \pi). \quad \rhd
\end{aligned}$$

Damos ahora la definición de diferenciabilidad de un campo escalar.

DEFINICIÓN. Sean, una función $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ y $p \in \Omega$ un punto con coordenadas (x, y, z). Se dice que f es **diferenciable** en p, si existe una función real de variable vectorial $g(h)$ tal que

$$\begin{cases} f(p+h) - f(p) = ||h||\, g(h) + < h, \nabla f_p > \\ \lim_{h \to (0,0,0)} g(h) = 0 \end{cases}$$

suponiendo que el vector gradiente del campo f,

$$\nabla f_p = \left(\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p), \frac{\partial f}{\partial z}(p) \right)$$

existe en el punto p.

Si la función $f : \Omega \to \mathbb{R}$ es diferenciable en cada punto de Ω, simplemente diremos que f es **diferenciable** en Ω.

A continuación damos una condición necesaria para que un campo escalar tenga la propiedad de diferenciabilidad. Su prueba se omite, pero el lector puede referirse a Lima (1992).

TEOREMA 4.2 (de diferenciabilidad). *Sea $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ y supóngase que en el punto $p \in \Omega$ las funciones determinadas por las derivadas parciales*

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} : \Omega \to \mathbb{R}$$

existen y son continuas (véase el capítulo 8 en los apéndices) en el punto p. entonces f es diferenciable en p.

Un corolario inmediato a este teorema es el siguiente.

COROLARIO 4.1 *Toda función $f : \Omega \to \mathbb{R}$ que involucre a las funciones clásicas en las variables independientes x, y, z, es una función diferenciable.*

EJEMPLOS. Las siguientes funciones son diferenciables en su dominio natural.

a. Cualquier polinomio $P(x, y, z)$ en las variables x, y, z, por ejemplo,

$$P(x, y, z) = 6xyz^3 + 2x^2yz - \frac{1}{2}xy^2z^2 + 3$$

b. $f(x, y, z) = \frac{x-y+z}{x^2+y^2+3z}$

c. $f(x, y, z) = \operatorname{sen} \sqrt{1 - e^{xy} - z}$



d. $f(x,y) = \ln\left(1 - x^{1/3} + \frac{1}{y^2}\right)$

e. $f(x,y,z) = \arctan(y/x + z)$

f. $f(x,y,z) = x^{y^2+\ln(z+x)}$

Un resultado inmediato cuya prueba tampoco realizaremos es el siguiente, referente, a la aritmética de las funciones diferenciables.

LEMA 4.2 *La suma (y diferencia), el producto y el cociente de funciones reales de variable vectorial que son diferenciables, es una función diferenciable*

EJEMPLO. Las siguientes funciones son diferenciables en todo su dominio natural.

a. $f(x,y,z) = 6xyz^2 + \frac{x-y+z}{x^2+y^2+3z}$

b. $f(x,y,z) = \operatorname{sen}\sqrt{1+e^{xy}-z} + \ln\left(1 - \sqrt[3]{xy} + \frac{z}{y^2}\right)$

c. $f(x,y,z) = \arctan\left(\frac{y}{x}\right) + x^{yz} + \ln(z+x)$

4.5 La regla de la cadena

En muchas ocasiones algunos problemas que involucran funciones de variable vectorial pueden reducirse a problemas de funciones reales de variable real que son más simples de estudiar, mediante la composición con curvas contenidas en las regiones donde están definidas tales funciones. A continuación damos el procedimiento.

Sean, $f : \Omega \to R$ una función diferenciable en todo punto de la región $\Omega \subset R^3$, y $\gamma : (a,b) \to \Omega$ una curva diferenciable. Entonces la composición

$$f \circ \gamma : (a,b) \to R$$

es una función real de variable real definida por,

$$(f \circ \gamma)(t) = f(\gamma(t))$$

EJEMPLO. Sean, el campo escalar,

$$f(x,y,z) = xe^{yz} + ye^{zx}$$

y la curva diferenciable

$$\gamma(t) = (4\cos t, 4\operatorname{sen} t,\ 2t).$$

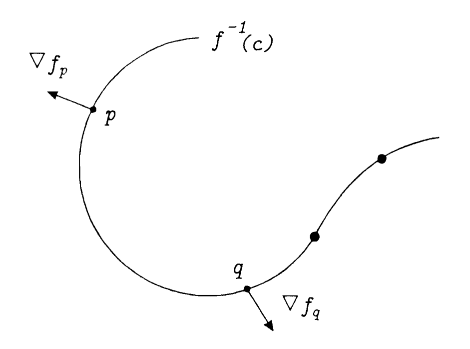


Figura 4.16: Gradiente de una función.

◁ Entonces

$$(f \circ \gamma)(t) = f(\gamma(t)) = f(4\cos t,\, 4\,\mathrm{sen} t,\, 2t)$$

$$= 4\cos t\, e^{(4\mathrm{sen} t)(2t)} + 4\mathrm{sen} t\, e^{(2t)(4\cos t)} = 4e^{8t\mathrm{sen} t}\cos t + 4e^{8t\cos t}\mathrm{sen} t$$

es decir,

$$(f \circ \gamma)(t) = 4e^{8t\mathrm{sen} t}\cos t + 4e^{8t\cos t}\mathrm{sen} t$$

resulta ser una función real de variable real, diferenciable.

EJEMPLO. Sean la función

$$f(x, y) = \ln(xy)$$

y la curva diferenciable $\gamma(t) = \left(t, \frac{1}{t^2}\right)$ con $t > 0$.

◁ Calculamos el dominio de f,

$$xy > 0, \quad \text{si y sólo sí,} \quad \begin{cases} x > 0 \quad \text{y} \quad y > 0 \\ \quad\quad\ \text{ó} \\ x < 0 \quad \text{y} \quad y < 0 \end{cases}$$

lo que nos dice que la unión de cuadrantes positivo (estricto) y negativo (estricto)

$$\Omega = \{(x, y) \mid (x > 0,\, y > 0),\, (x < 0,\, y < 0)\}$$

es el dominio de f, y que la curva

$$\gamma(t) = (t,\, 1/t^2), \quad t > 0$$

está contenida en Ω.

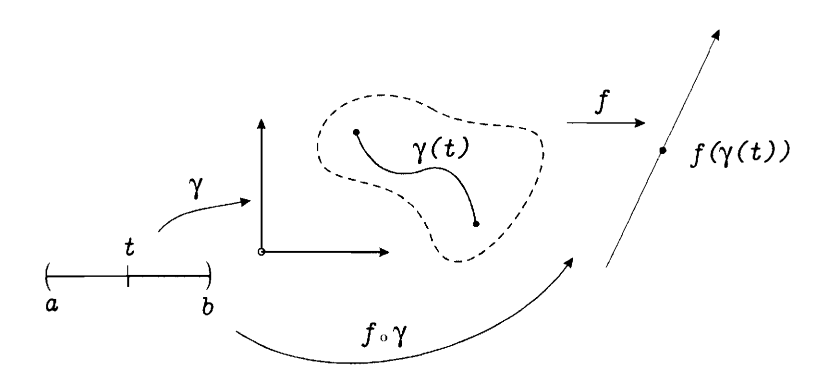


Figura 4.17: Regla de la cadena.

Entonces, la composición $(f\circ\gamma)(t)$, que resulta ser una función diferenciable, se calcula por

$$(f\circ\gamma)(t) = f(\gamma(t)) = f(t, 1/t^2) = \ln(t\,1/t^2) = \ln(1/t) = -\ln t \quad \triangleright$$

Damos una fórmula para calcular la derivada de una composición de funciones como la que se ha mencionado (véase la figura 4.17).

TEOREMA 4.3 (Regla de la cadena) *Para una curva suave $\gamma(t)$ y una función diferenciable $f(x, y, z)$ se cumple la fórmula*

$$\frac{d(f\circ\gamma)}{dt}(t) = < \nabla f_{\gamma(t)}, \dot{\gamma}(t) >$$

que es la derivada de la composición $f\circ\gamma$ en el tiempo t.

Esto es, en coordenadas x, y, z del espacio se tiene

$$\frac{d(f\circ\gamma)}{dt}(t) = \left\langle \left(\frac{\partial f}{\partial x}(\gamma(t)), \frac{\partial f}{\partial y}(\gamma(t)), \frac{\partial f}{\partial z}(\gamma(t))\right), \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right)\right\rangle$$

$$= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial x}\frac{dz}{dt}$$

$\triangleleft$ Sea $t \in (a, b)$ un punto arbitrario y considérese el incremento vectorial

$$h = \gamma(t+s) - \gamma(t)$$

Es claro que $h \to (0, 0, 0)$ si $s \to 0$.

En virtud de que f es diferenciable, particularmente en el punto vectorial $\gamma(t)$, se cumple que

$$f(\gamma(t)+h) - f(\gamma(t)) =< \nabla f_{\gamma(t)}, h > + ||h|| g(h)$$

donde $\lim_{h\to(0,0,0)} g(h) = 0$, siendo h cualquier incremento vectorial.

Si se toma $h = \gamma(t+s) - \gamma(t)$, entonces se tiene

$$f(\gamma(t+s)) - f(\gamma(t)) =< \nabla f_{\gamma(t)}, \gamma(t+s)-\gamma(t) > + ||\gamma(t+s)-\gamma(t)|| g(h)$$

que al dividir entre el incremento s nos da

$$\frac{(\gamma(t+s) - f(\gamma(t))}{s} = \left\langle \nabla f_{\gamma(t)}, \frac{\gamma(t+s)-\gamma(t)}{s} \right\rangle + \left|\left| \frac{\gamma(t+s)-\gamma(t)}{s} \right|\right| g(h)$$

Si tomamos el límite cuando $s \to 0$ se tiene que

$$\frac{d(f\circ\gamma)}{dt}(t) = \lim_{s\to 0} \frac{f(\gamma(t+s)) - f(\gamma(t))}{s} = \left\langle \nabla f_{\gamma(t)}, \lim_{s\to 0} \frac{\gamma(t+s)-\gamma(t)}{s} \right\rangle$$

$$+ \left|\left| \lim_{s\to 0} \frac{\gamma(t+s)-\gamma(t)}{s} \right|\right| \lim_{h\to 0} g(h) = \langle \nabla f_{\gamma(t)}, \dot\gamma(t) \rangle + ||\dot\gamma(t)|| \lim_{h\to 0} g(h)$$

$$=< \nabla f_{\gamma(t)}, \dot\gamma(t) >$$

debido a que $\lim_{h\to(0,0,0)} g(h) = 0$. ▷

EJEMPLO. Sea la función diferenciable

$$f(x,y) = x^2 + 2xy$$

y sean $x = r\cos\varphi$, $y = r\,\mathrm{sen}\,\varphi$ con $\varphi \in [0, 2\pi]$.

◁ Al formar la función $g(r,\varphi) = f(r\cos\varphi, r\,\mathrm{sen}\,\varphi)$ se tiene, al derivar la función g respecto a su variable verdadera r,

$$\frac{\partial g}{\partial r} = \frac{\partial f}{\partial x}\frac{dx}{dr} + \frac{\partial f}{\partial y}\frac{dy}{dr} = (2x+2y)(\cos\varphi) + (2y)(\mathrm{sen}\varphi)$$

$$= 2r\cos^2\varphi + 2r\mathrm{sen}\varphi\cos\varphi + 2r\cos\varphi\,\mathrm{sen}\varphi = 2r\cos^2\varphi + 4\,\mathrm{sen}\varphi\cos\varphi$$

Si φ es variable también, entonces, al derivar la función g respecto a su variable explícita φ, se tiene que

$$\frac{\partial g}{\partial\varphi} = \frac{\partial f}{\partial x}\frac{dx}{d\varphi} + \frac{\partial f}{\partial y}\frac{dy}{d\varphi}$$

$$= (2x+2y)(-r\mathrm{sen}\,\varphi) + (2x)(r\cos\varphi)$$



$$= (2r\cos\varphi + 2r\,\mathrm{sen}\,\varphi)(-r\,\mathrm{sen}\,\varphi) + 2r\cos\varphi\, r\cos\varphi$$
$$= -2r^2\mathrm{sen}\,\varphi\cos\varphi - 2r^2\mathrm{sen}^2\varphi + 2r^2\cos^2\varphi \quad \rhd$$

EJEMPLO. Considere el campo escalar de tres variables

$$w = f(x, y, z) = e^{xy}\cos z$$

y las funciones

$$\begin{cases} x = tu \\ y = \mathrm{sen}(tu) \\ z = u^2 \end{cases}$$

$\lhd$ Al calcular la variación de w respecto a t se tiene,

$$\frac{\partial w}{dt} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t}$$
$$= (ye^{xy}\cos z)u + (xe^{xy}\cos z)u\cos(tu) - (e^{xy}\mathrm{sen}\,z)(0)$$
$$= ue^{xy}\cos z[y + x\cos(tu)] = ue^{tu\,\mathrm{sen}\,(tu)}\cos u^2[\mathrm{sen}\,tu + tu\cos(tu)] \quad \rhd$$

EJEMPLO. Dada la función

$$u = f(x^2 - y,\, xy)$$

calcular las derivadas parciales,

$$\frac{\partial u}{\partial x}, \quad \frac{\partial u}{\partial y}$$

$\lhd$ Construimos variables explícitas para la función f, dadas por

$$\begin{cases} r = x^2 - y \\ q = xy \end{cases}$$

Tenemos entonces que al derivar la función u respecto a su variable explícita x,

$$\frac{\partial u}{\partial x} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial r}{\partial q}\frac{\partial q}{\partial x} = \frac{\partial f}{\partial r}(2x) + \frac{\partial f}{\partial q}(y) = 2x\frac{\partial f}{\partial r} + y\frac{\partial f}{\partial q}$$

Por otro lado, cuando se deriva u respecto a su variable explícita y se obtiene

$$\frac{\partial u}{\partial y} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial f}{\partial q}\frac{\partial q}{\partial y} = \frac{\partial f}{\partial r}(-1) + \frac{\partial f}{\partial q}(x) = x\frac{\partial f}{\partial q} - \frac{\partial f}{\partial r} \quad \rhd$$

EJEMPLO. Sea la función

$$g(t, x, y) = f(t^2x,\, ty)$$

Calcular las derivadas parciales

$$\frac{\partial g}{\partial t}, \quad \frac{\partial g}{\partial x}, \quad \frac{\partial g}{\partial y}$$

◁ Entendemos por $u = t^2x$, $v = ty$ a las coordenadas explícitas de f. Entonces,

$$\frac{\partial g}{\partial t} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial t} = \left(\frac{\partial f}{\partial u}\right)(2tx)\left(\frac{\partial f}{\partial v}\right)(y)$$

$$= 2tx\,\frac{\partial f}{\partial u} + y\frac{\partial f}{\partial v}$$

Por otro lado,

$$\frac{\partial g}{\partial x} = \frac{\partial f}{\partial u}\frac{du}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} = \frac{\partial f}{\partial u}(t^2) + \frac{\partial f}{\partial v}(0) = t^2\,\frac{\partial f}{\partial u}$$

Un cálculo análogo prueba que,

$$\frac{\partial g}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} = t\,\frac{\partial f}{\partial v} \quad ▷$$

EJEMPLO. Supóngase que en una habitación $\Omega \subset \mathbb{R}^3$ está definida una función $f : \Omega \to \mathbb{R}$ tal que en cada punto p de Ω nos da una temperatura.

◁ Si un insecto recorre una trayectoria en Ω tal que en cada punto de la trayectoria la temperatura f es constante c, al conocer tal trayectoria por $\gamma(t)$ con $t \in [a, b]$, se tiene que

$$f(\gamma(t)) = c$$

Al derivar ambos lados esta ecuación se tiene,

$$\frac{d}{dt}f(\gamma(t)) = 0$$

pero, por la regla de la cadena esto implica que

$$\langle \nabla f_{\gamma(t)}, \dot{\gamma}(t) \rangle = 0$$

Lo anterior nos dice que los vectores gradiente $\nabla f_{\gamma(t)}$ y velocidad $\dot{\gamma}(t)$ son ortogonales para todo punto $\gamma(t)$ de la trayectoria (véase la figura 4.18).

Se dice en este caso, que el insecto realizó una trayectoria con la misma temperatura (isotermal). ▷

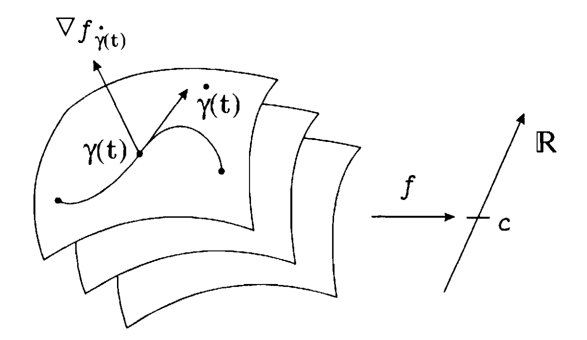


Figura 4.18: Trayectoria isotermal.

4.6 Derivada direccional

Introducimos ahora un concepto fundamental a la hora de hacer el análisis de un **campo escalar** (función real de variable vectorial) definido en una región $\Omega \subset \mathbb{R}^3$.

Sea $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ una función diferenciable y, sean, el punto $p \in \Omega$ y ξ un vector unitario.

La línea recta que pasa por p y con dirección ξ tiene una ecuación paramétrica dada por

$$\gamma(t) = p + t\xi \quad \text{con} \quad \gamma(0) = p,\ \dot{\gamma}(0) = \xi$$

Para los puntos sobre el segmento de recta, contenido en Ω, se tiene un valor real bajo la función f dado por

$$f(\gamma(t)) = f(p + t\xi)$$

Si calculamos la **variación** de f a lo largo de tal línea obtenemos, por la regla de la cadena,

$$\frac{d(f \circ \gamma)}{dt}(t) = \left\langle \nabla f_{\gamma(t)}, \dot{\gamma}(t) \right\rangle$$

Particularmente, en $t = 0$ se tiene

$$\frac{df(p + t\xi)}{dt}|_{t=0} = \left\langle \nabla f_{\gamma(0)}, \dot{\gamma}(0) \right\rangle = (\nabla f_p, \xi\rangle$$

DEFINICIÓN. Tal variación de f se llamará la **derivada direccional** de f en el punto $p \in \Omega$ en la dirección de ξ.

En coordenadas x, y, z se tiene la igualdad

$$\frac{df}{dt}(p+t\xi)|_{t=0} = \left\langle \left(\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p), \frac{\partial f}{\partial z}(p) \right), (\xi^1, \xi^2, \xi) \right\rangle$$

$$= \frac{\partial f}{\partial x}(p)\xi^1 + \frac{\partial f}{\partial y}(p)\xi^2 + \frac{\partial f}{\partial z}(p)\xi^3$$

Observamos que para cuando se toma un vector localizado arbitrario $\eta \neq 0$ con extremo inicial en p, también se puede hablar de la derivada direccional de la función f en dirección η con el simple hecho de calcular la derivada de f en la dirección del normalizado $\xi = \frac{\eta}{||\eta||}$, en virtud de que ξ y η tienen la misma dirección y sentido (véase la figura 4.19).

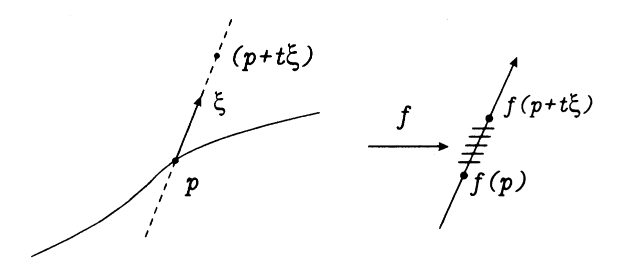


Figura 4.19: Derivada direccional de f en p en la dirección ξ.

Una notación apropiada para conocer a la derivada direccional de la función f en el punto p, en la dirección del vector ξ es,

$$D_\xi f_p = \frac{df}{dt}(p+t\xi)\,|_{t=0}$$

EJEMPLO. Sea el campo escalar

$$f(x,y) = x^2 + y^2$$

Calcular la derivada direccional de f en el punto $p = (-1, 3)$ en la dirección $\eta = (1, 2)$

$\lhd$ Como el vector η no es unitario pues $||\eta|| = \sqrt{5}$, hemos de hacerlo unitario mediante la cadena de igualdades,

$$\xi = \frac{\eta}{||\eta||} = \frac{(1,2)}{\sqrt{5}} = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right)$$

De esta manera,

$$D_\xi\, f_p = \langle \nabla f_p, \xi \rangle = \left\langle (2x, 2y)\,|_{(-1,3)}, \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right) \right\rangle$$



$$= \left\langle (-2,6), \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)\right\rangle = \frac{-2}{\sqrt{5}} + \frac{12}{\sqrt{5}} = \frac{10}{\sqrt{5}} = 2\sqrt{5} \quad \triangleright$$

OBSERVACIÓN. Si en particular se toman el vector canónico $\xi = e_1 = (1,0,0)$ se tiene que

$$D_{e_1} f_p = \left\langle \left(\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p), \frac{\partial f}{\partial z}(p)\right), (1,0,0)\right\rangle = \frac{\partial f}{\partial x}(p)$$

que es la derivada parcial ordinaria respecto a la primer coordenada.

Análogamente,

$$D_{e_2} f_p = \frac{\partial f}{\partial y}(p) \quad \text{y} \quad D_{e_3} f_p = \frac{\partial f}{\partial z}(p)$$

OBSERVACIÓN. De la cadena de igualdades

$$D_\xi f_p = \langle \nabla f_p, \xi \rangle = ||\nabla f_p|| \, ||\xi|| \cos\theta = ||\nabla f_p|| \cos\theta$$

donde θ es el ángulo formado por ∇f_p y el vector unitario ξ, se tiene una función en la variable θ definida por,

$$h(\theta) = ||\nabla f|| \, \cos\theta$$

Sabiendo que la relación de ξ con θ está determinada de manera biunívoca, tenemos que

a. h tiene un valor máximo en $\theta = 0$, es decir, cuando ξ está en el mismo sentido y dirección que ∇f_p.

b. h tiene un valor mínimo en $\theta = \pi$, *y* se obtiene es decir, cuando ξ está en dirección de ∇f_p pero con sentido contrario.

De esta manera, *la función f crece más rápidamente alrededor del punto p en dirección del vector gradiente ∇f_p, y decrece más rápidamente en dirección del vector $-\nabla f_p$.*

EJEMPLO. Sea el campo escalar

$$f(x,y) = x^2 + xy + y^2$$

$\triangleleft$ De la anterior discusión, en el punto $p = (-1,1)$ la función crece más rápidamente en dirección del vector

$$\eta = \nabla f_p = (2x + y, \, 2y + x)_{(-1,1)} = (-1,1)$$



Al normalizar η obtenemos el vector unitario

$$\xi = \frac{\eta}{||\eta||} = \frac{(-1,1)}{\sqrt{2}} = \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

que nos permite calcular la derivada direccional en dirección de η, mediante

$$D_\xi f_p =< \nabla f_p, \xi >= \left\langle (-1,1), \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right\rangle = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} = \frac{2}{\sqrt{2}} = \sqrt{2}$$

Por otro lado, la función f decrece más rápidamente en p en la dirección del vector $-\eta = -\nabla f_p = (1,-1)$. La derivada direccional de la función f en el punto p, en tal dirección es, en este caso,

$$D_{-\xi} f_p =< \nabla f_p, -\xi >= \left\langle (-1,1), \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)\right\rangle = -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} = -\sqrt{2} \quad \triangleright$$

4.7 El Teorema de Taylor

En esta parte generalizamos El **teorema** de **Taylor**, para funciones reales con variable vectorial que están definidas en regiones del espacio euclidiano. Lo hacemos apenas para funciones suaves de dos variables, suponiendo una extensión natural para el caso de más variables.

La idea básica del teorema consiste en aproximar una función de cierta clase de diferenciabilidad por un polinomio en el número de variables independientes, de la misma forma que para el caso de una sola variable (véase **Reyes**, 1996).

Comenzamos la sección con el concepto de derivada **parcial de orden superior**.

EJEMPLO. Sea el campo escalar diferenciable

$$f(x,y) = e^{x^2+y^2}$$

$\triangleleft$ Al calcular sus derivadas parciales se obtienen las nuevas funciones

$$\frac{\partial f}{\partial x}(x,y) = 2xe^{x^2+y^2} = g_1(x,y)$$

$$\frac{\partial f}{\partial y}(x,y) = 2y\, e^{x^2+y^2} = g_2(x,y)$$

Claramente g_1 y g_2 también son funciones diferenciables y se puede calcular sus derivadas parciales

$$\frac{\partial g_1}{\partial x}(x,y) = 2e^{x^2+y^2} + 2x\,2xe^{x^2+y^2} = (2+4x^2)e^{x^2+y^2}$$

$$\frac{\partial g_1}{\partial y}(x,y) = 2y2xe^{x^2+y^2} = 4xye^{x^2+y^2}$$

$$\frac{\partial g_2}{\partial x}(x,y) = 2x2ye^{x^2+y^2} = 4xye^{x^2+y^2}$$

$$\frac{\partial g_2}{\partial y}(x,y) = 2e^{x^2+y^2} + 2y\,2ye^{x^2+y^2} = (2+4y^2)e^{x^2+y^2}$$

que resultan ser nuevamente funciones diferenciables.

Al omitir (por notación) las funciones g_1, g_2 de las ecuaciones anteriores, éstas se pueden escribir,

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = (2+4x^2)e^{x^2+y^2}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = 4xye^{x^2+y^2}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = 4xye^{x^2+y^2}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = (2+4y^2)e^{x^2+y^2}$$

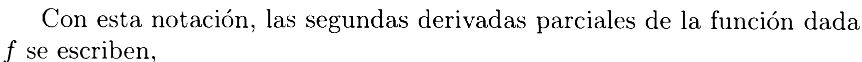
De ahora y en adelante, adoptaremos la siguiente notación para las llamadas **segundas derivadas parciales** de la función f,

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x \partial x} = \frac{\partial^2 f}{\partial x^2}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y \partial y} = \frac{\partial^2 f}{\partial y^2}$$



Con esta notación, las segundas derivadas parciales de la función dada f se escriben,

$$\frac{\partial^2 f}{\partial x^2} = (2 + 4x^2)e^{x^2+y^2}$$

$$\frac{\partial^2 f}{\partial y \partial x} = 4xye^{x^2+y^2}$$

$$\frac{\partial^2 f}{\partial x \partial y} = 4xye^{x^2+y^2}$$

$$\frac{\partial^2 f}{\partial y^2} = (2 + 4y^2)e^{x^2+y^2}$$

Observamos que para este caso se cumple la igualdad de las relaciones obtenidas para las segundas **derivadas parciales mixtas**:

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} = 4xye^{x^2+y^2} \quad \triangleright$$

DEFINICIÓN. Si la función $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ es tal que sus derivadas parciales existen

$$\frac{\partial f}{\partial x} : \Omega \to \mathbb{R}$$

$$\frac{\partial f}{\partial y} : \Omega \to \mathbb{R}$$

y como funciones definidas en Ω son continuas, se dice que la función es de **clase C^1**. Si además las funciones $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ son diferenciables y sus derivadas parciales (segundas derivadas parciales de la función f)

$$\frac{\partial^2 f}{\partial x^2} : \Omega \to \mathbb{R}$$

$$\frac{\partial^2 f}{\partial y \partial x} : \Omega \to \mathbb{R}$$

$$\frac{\partial^2 f}{\partial x \partial y} : \Omega \to \mathbb{R}$$

$$\frac{\partial^2 f}{\partial y^2} : \Omega \to \mathbb{R}$$

son continuas, entonces f se dirá una **función de clase C^2**.

Las derivadas parciales segundas

$$\frac{\partial^2 f}{\partial y \partial x}, \frac{\partial^2 f}{\partial x \partial y}$$

se llaman las segundas **derivadas parciales mixtas**.

En el ejemplo anterior se indicó una propiedad de igualdad de tales parciales mixtas. Una condición suficiente para la igualdad en general se tiene el teorema siguiente, cuya prueba es muy técnica y la omitimos (véase **Lima**, 1992).

TEOREMA 4.4 (Conmutatividad de las segundas parciales mixtas). *Sea $f : \Omega \rightarrow \mathbb{R}$ campo escalar de clase C^2, esto es, las funciones obtenidas para las segundas derivadas parciales son continuas (véase el capítulo 8), entonces se cumple la igualdad de conmutatividad,*

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial x}$$

EJEMPLO. Dada la función

$$f(x, y) = x^2 y^3 + 3xy$$

mostramos la validez del teorema 4.4 para este campo escalar.

◁ Al calcular la derivada parcial de la función respecto a x se tiene

$$\frac{\partial f}{\partial x} = 2xy^3 + 3y$$

y consecuentemente, al derivar esta última función respecto a y obtenemos

$$\frac{\partial^2 f}{\partial y \partial x} = 6xy^2 + 3$$

Por otro lado,

$$\frac{\partial f}{\partial y} = 3x^2 y^2 + 3x$$

lo que implica

$$\frac{\partial^2 f}{\partial x \partial y} = 6xy^2 + 3 = \frac{\partial^2 f}{\partial x \partial y} \quad \triangleright$$

Inductivamente, podemos definir a una función $f(x, y)$ de clase C^r (con $r \geq 1$), si para cualquier partición del entero r en los enteros s y m, la derivada parcial r−ésima de f dada por

$$\frac{\partial^r f}{\partial y^m \partial x^s} : \Omega \rightarrow R$$

existe y es una función continua. Aquí, $0 \leq s \leq r$, $0 \leq m \leq r$ y $r = m + s$.



En la expresión

$$\frac{\partial^r f}{\partial y^m \partial x^s}$$

se entiende que se calcula $m-$derivadas parciales de f respecto a la variable y y $s-$derivadas parciales respecto a x, realizando en total un cálculo de $m + s = r$ derivadas parciales.

La extensión de la definición de clase de diferenciabilidad para un campo escalar de más variables es una analogía a la dada para apenas dos variables.

El Teorema 4.4 se puede extender para una función que depende de más variables. Esto es, si $f(x, y, z)$ es una función que satisface las condiciones dadas de diferenciabilidad, el teorema de conmutatividad se cumple también.

De hecho, por ejemplo,

$$\frac{\partial^2 f}{\partial x \partial z} = \frac{\partial^2 f}{\partial z \partial x}$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

$$\frac{\partial^2 f}{\partial z \partial y} = \frac{\partial^2 f}{\partial y \partial z}$$

Si la misma función $f(x, y, z)$ es de clase C^3, entonces la conmutatividad se extiende de manera natural.

EJEMPLO. Consideremos a $f(x, y, z)$ un campo escalar función de clase C^3 definido en alguna región del espacio.

$\lhd$ Se tiene la siguiente cadena de igualdades,

$$\frac{\partial^3 f}{\partial y \partial x^2} = \frac{\partial^2}{\partial y \partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2}{\partial x \partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y \partial x}\right)$$

$$= \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial^3 f}{\partial x^2 \partial y}$$

También se cumplen las igualdades,

$$\frac{\partial^3 f}{\partial z \partial y \partial x} = \frac{\partial^3 f}{\partial z \partial x \partial y} = \frac{\partial^3 f}{\partial x \partial y \partial z} = \frac{\partial^3 f}{\partial x \partial z \partial y}$$

así como la igualdad,

$$\frac{\partial^3 f}{\partial y \partial z^2} = \frac{\partial^3 f}{\partial z^2 \partial y} \quad \rhd$$

Si $f(x, y, z)$ es una función de clase C^4, también se cumple el teorema de conmutatividad de las derivadas parciales. Por ejemplo, son justas las igualdades

$$\frac{\partial^4 f}{\partial z \partial y \partial x^2} = \frac{\partial^4 f}{\partial x^2 \partial z \partial y} = \frac{\partial^4 f}{\partial y \partial x^2 \partial z}$$

$$\frac{\partial^4 f}{\partial x \partial z^3} = \frac{\partial^4 f}{\partial z^3 \partial x} = \frac{\partial^4 f}{\partial z \partial x \partial z^2}$$

$$\frac{\partial^4 f}{\partial y^2 \partial x^2} = \frac{\partial^4 f}{\partial x^2 \partial y^2} = \frac{\partial^4 f}{\partial y \partial x \partial y \partial x}$$

Generalizando lo anterior, si $f(x, y, z)$ es una función de clase C^r con $(r \geq 1)$, el teorema de conmutatividad también se extiende. Por ejemplo,

$$\frac{\partial^r f}{\partial z^\ell \partial y^m \partial x^n} = \frac{\partial^r f}{\partial y^m \partial x^n \partial z^\ell} = \frac{\partial^r f}{\partial x^n \partial z^\ell \partial y^m}$$

donde $\ell + m + n = r$. Aquí se entiende que la función f se va a derivar parcialmente r veces en total: ℓ veces respecto a la variable z, m veces respecto a la variable y y m veces respecto a la variable x.

EJEMPLO. Calcular las derivadas parciales

$$\frac{\partial^3 f}{\partial x \partial y \partial z}, \frac{\partial^3 f}{\partial z \partial y \partial x}$$

para la función

$$f(x, y, z) = e^{xyz}$$

$\lhd$ En la notación dada, tenemos que

$$\frac{\partial f}{\partial z} = xye^{xyz}$$

$$\frac{\partial^2 f}{\partial y \partial z} = xe^{xyz} + x^2yze^{xyz} = (x + x^2yz)e^{xyz}$$

de donde,

$$\frac{\partial^3 f}{\partial x \partial y \partial z} = (1 + 2xyz)e^{xyz} + (x + x^2yz)yze^{xyz}$$

$$= (1 + 3xyz + x^2y^2z^2)e^{xyz}$$

Por otro lado,

$$\frac{\partial f}{\partial x} = yze^{xyz}$$

$$\frac{\partial^2 f}{\partial y \partial x} = ze^{xyz} + yz^2xe^{xyz} = (z + yxz^2)e^{xyz}$$

$$\frac{\partial^3 f}{\partial z \partial y \partial x} = (1 + 2xyz)e^{xyz} + (z + yxz^2)xye^{xyz}$$

$$= (1 + 3xyz + x^2y^2z^2)e^{xyz}$$

lo que verifica las igualdades de las **parciales mixtas**. ▷

Ahora construimos el polinomio de Taylor de una función suave con dos variables.

Recordamos que para funciones reales de variable real se tenía que si la función $f : (a, b) \to \mathbb{R}$ es de clase C^r, entonces, por El teorema de Taylor, alrededor del punto $p \in (a, b)$ la función f se escribe localmente

$$f(x) = f(p) + \frac{f'(p)}{1!}(x - p) + \frac{f''(p)}{2!}(x - p)^2 + \frac{f'''(p)}{3!}(x - p)^3$$

$$+ \cdots + \frac{f^{(r)}(p)}{r!}(x - p)^r + o((x - p)^{r+1})$$

donde $o((x - p)^{r+1})$ significa un error en el monomio $x - p$ de orden $r + 1$ (véase **Reyes**, 1996).

Aquí, la expresión polinomial de grado r en el monomio $x - p$

$$f(p) + \frac{f'(p)}{1!}(x - p) + \frac{f''(p)}{2!}(x - p)^2 + \frac{f'''(p)}{3!}(x - p)^3 + \cdots + \frac{f^{(r)}(p)}{r!}(x - p)^r$$

se llama El **polinomio de Taylor** de grado r de la función f en el punto p.

Generalizamos la idea de aproximar una función en dos variables $f(x, y)$ mediante un polinomio de grado r en dos variables x, y, para cuando la función dada es de clase C^r.

Un **polinomio lineal** en las variables x, y es del tipo

$$P_1(x, y) = c_{00} + c_{10}(x - a) + c_{01}(y - b)$$

donde, los coeficientes c_{ij} son constantes reales, al igual que a y b.

Un **polinomio cuadrático** en las variables x, y es del tipo:

$$P_2(x, y) = c_{00} + c_{10}(x - a) + c_{01}(y - b)$$

$$+ c_{20}(x - a)^2 + c_{11}(x - a)(y - b) + c_{02}(y - b)^2$$

donde, nuevamente, los coeficientes c_{ij} son constantes reales.



Un **polinomio cúbico** tiene la forma

$$P_3(x,y) = c_{00} + \underbrace{c_{10}(x-a) + c_{01}(y-b)}_{\text{parte lineal}}$$

$$+ \underbrace{c_{20}(x-a)^2 + c_{11}(x-a)(y-b) + c_{02}(y-b)^2}_{\text{parte cuadrática}}$$

$$+ \underbrace{c_{30}(x-a)^3 + c_{21}(x-a)^2(y-b) + c_{12}(x-a)(y-b)^2 + c_{03}(y-b)^3}_{\text{parte cúbica}}$$

Observamos que, en general, un polinomio de grado r en las variables x, y se escribe,

$$P_r(x,y) = \sum_{k=1}^{r} \left(\sum_{k=i+j} c_{ij}(x-a)^i\,(y-b)^j \right) + c_{00}$$

Iniciamos un proceso de aproximación de cualquier función de dos variables que sea suave, por un polinomio en dos variables.

Sea $f(x,y)$ una función arbitraria de clase C^r definida en una región Ω, en el punto $p = (a,b) \in \Omega$ y sea $(x,y) \in \Omega$ cualquier punto cercano a p.

Primeramente intentamos aproximar a f linealmente en una cercanía del punto p, mediante el polinomio

$$f(x,y) = f(p) + \underbrace{c_{10}(x-a) + c_{01}(x-b)}_{\text{parte lineal}} + o((x-a)^2, (y-b)^2)$$

donde la expresión $o((x-a)^2, (x-b)^2)$ define un **error** en los monomios $x-a$ y $y-b$ de orden dos.

◁ Calculamos los coeficientes c_{ij} apropiados, calculando primeramente

$$\frac{\partial f}{\partial x} = c_{10} + o((x-a), (y-b)^2),$$

lo que nos dice que, al calcular en $p = (a,b)$, esto es, en $x = a$, $y = b$,

$$\frac{\partial f}{\partial x}(a,b) = c_{10} + o((x-a), (y-b))|_{x=a,\,y=b} = c_{10}$$

lo que nos da el primer coeficiente.



Igualmente, de la relación

$$\frac{\partial f}{\partial y} = c_{01} + o((x-a)^2, (y-b))$$

se obtiene, al evaluar en $p = (a, b)$,

$$\frac{\partial f}{\partial y}(a,b) = c_{01}$$

Por lo tanto, la aproximación lineal de f alrededor de p sería mediante la expresión

$$f(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) + o((x-a)^2, (y-b)^2) \quad \triangleright$$

Aproximamos ahora a la función $f(x, y)$ hasta el orden dos, mediante el polinomio cuadrático

$$f(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b)$$

$$\underbrace{+c_{20}(x-a)^2 + c_{11}(x-a)(y-b) + c_{02}(y-b)^2}_{\text{parte cuadrática}} + o((x-a)^3, (y-b)^3)$$

donde la expresión $o((x-a)^3, (y-b)^3)$ define un **error** de orden cúbico en los monomios indicados.

$\triangleleft$ Procedemos a calcular las constantes c_{20}, c_{11}, c_{02} apropiadas del polinomio.

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}(a,b) + 2c_{20}(x-a) + c_{11}(y-b) + o((x-a)^2, (y-b)^3)$$

donde la expresión $o((x-a)^2, (y-b)^3)$ define un error de orden dos en el monomio $x - a$ junto con un error de orden tres en el monomio $y - b$.

Por lo tanto,

$$\frac{\partial^2 f}{\partial x^2} = 2c_{20} + o((x-a), (y-b)^3)$$

lo que al evaluar en el punto (a, b) nos dice que

$$\frac{\partial^2 f}{\partial x^2}(a,b) = 2c_{20}$$

es decir,

$$c_{20} = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(a,b)$$



Además, se tiene

$$\frac{\partial^2 f}{\partial y \partial x} = c_{11} + o((x-a)^2, (y-b)^2)$$

lo que nos indica, al elvaluar en el punto (a, b),

$$\frac{\partial^2 f}{\partial y \partial x}(a,b) = c_{11}$$

Por otro lado, al derivar parcialmente respecto a y la función inicial f se tiene

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial y}(a,b) + c_{11}(x-a) + 2c_{02}(y-b) + o((x-a)^3, (y-b)^2)$$

con lo que,

$$\frac{\partial^2 f}{\partial y^2} = 2c_{02} + o((x-a)^3, (y-b))$$

Al evaluar en el punto $p = (a, b)$ esta expresión se obtiene,

$$\frac{\partial^2 f}{\partial y^2}(a,b) = 2c_{02}$$

y por lo tanto,

$$c_{02} = \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(a,b)$$

De esta forma, hasta el orden dos la aproximación de la función f es,

$$f(x,y) = f(p) + \frac{\partial f}{\partial x}(p)(x-a) + \frac{\partial f}{\partial y}(p)(y-b)$$

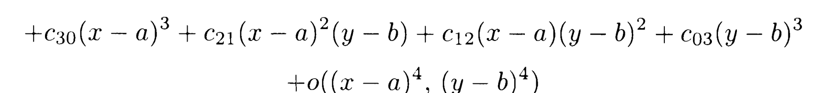
$$+\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(p)(x-a)^2 + \frac{\partial^2 f}{\partial y \partial x}(p)(x-a)(y-b) + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(p)(y-b)^2$$

$$+o((x-a)^3, (y-b)^3) \quad \triangleright$$

Un intento de aproximación a orden tres para f sería dada por el polinomio,

$$f(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b)$$

$$+\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(a,b)(x-a)^2 + \frac{\partial^2 f}{\partial y \partial x}(a,b)(x-a)(y-b) + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(a,b)(y-b)^2$$



$$+c_{30}(x-a)^3 + c_{21}(x-a)^2(y-b) + c_{12}(x-a)(y-b)^2 + c_{03}(y-b)^3$$
$$+o((x-a)^4,\ (y-b)^4)$$

donde $o((x-a)^4, (y-b)^4)$ es un error de orden cuatro en los monomios dados.

Calculamos los coeficientes $c_{30}, c_{21}, c_{12}, c_{03}$ de tal expresión utilizando el mismo procedimiento.

◁ Derivando parcialmente la igualdad respecto a la variable x, tenemos

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}(a,b) + \frac{\partial^2 f}{\partial x^2}(a,b)(x-a) + \frac{\partial^2 f}{\partial y \partial x}(a,b)(y-b)$$

$$+3c_{30}(x-a)^2 + 2c_{21}(x-a)(y-b) + c_{12}(y-b)^2 + o((x-a)^3, (y-b)^4)$$

de donde,

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial x^2}(a,b) + 3!\, c_{30}(x-a) + 2c_{21}(y-b) + o((x-a)^2, (y-b)^4)$$

y entonces,

$$\frac{\partial^3 f}{\partial x^3} = 3! c_{30} + o((x-a), (y-b)^4)$$

lo que nos indica, al evaluar el punto (a,b)

$$\frac{\partial^3 f}{\partial x^3}(a,b) = 3! c_{30}$$

Así, podemos calcular c_{30} por

$$c_{30} = \frac{\frac{\partial^3 f}{\partial x^3}(a,b)}{3!}$$

Por otro lado,

$$\frac{\partial^3 f}{\partial y \partial x^2} = 2c_{21} + o((x-a)^2, (y-b)^3)$$

de donde,

$$\frac{\partial^3 f}{\partial y \partial x^2}(a,b) = 2c_{21}$$

lo que implica

$$c_{21} = \frac{\frac{\partial^3 f}{\partial y \partial x^2}(a,b)}{2}$$



Además, de $\frac{\partial f}{\partial x}$ podemos obtener

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial y \partial x}(a,b) + 2c_{21}(x-a) + 2c_{12}(y-b) + o((x-a)^3,(y-b)^3)$$

y de esto,

$$\frac{\partial^3 f}{\partial y^2 \partial x} = 2c_{12} + o((x-a)^3,(y-b)^2)$$

Por lo tanto,

$$\frac{\partial^3 f}{\partial y^2 \partial x}(a,b) = 2c_{12}$$

con lo que se obtiene el coeficiente

$$c_{12} = \frac{\frac{\partial^3 f}{\partial y^2 \partial x}(a,b)}{2}$$

Otro cálculo análogo prueba que,

$$\frac{\partial^3 f}{\partial y^3}(a,b) = 3!c_{03}$$

es decir,

$$c_{03} = \frac{\frac{\partial^3 f}{\partial y^3}(a,b)}{3!}$$

Por lo tanto, el desarrollo a orden tres nos quedaría, al sustituir tales coeficientes,

$$f(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b)$$

$$+\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(a,b)(x-a)^2 + \frac{\partial^2 f}{\partial y \partial x}(a,b)(x-a)(y-b) + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(a,b)(y-b)^2$$

$$+\frac{\frac{\partial^3 f}{\partial x^3}(a,b)}{3!}(x-a)^3 + \frac{\frac{\partial^3 f}{\partial y \partial x^2}(a,b)}{2}(x-a)^2(y-b) + \frac{\frac{\partial^3 f}{\partial y^2 \partial x}(a,b)}{2}(x-a)(y-b)^2$$

$$+\frac{\frac{\partial^3 f}{\partial y^3}(a,b)}{3!}(y-b)^3 + o((x-a)^4,(y-b^4)) \quad \triangleright$$

Observamos que la relación anterior se puede escribir, al tomar $p = (a,b)$, también como

$$f(x,y) = f(p) + \frac{\frac{\partial f}{\partial x}(p)}{1!0!}(x-a) + \frac{\frac{\partial f}{\partial y}(p)}{0!1!}(y-b)$$

$$+\frac{\frac{\partial^2 f}{\partial x^2}(p)}{2!0!}(x-a)^2+\frac{\frac{\partial^2 f}{\partial y\partial x}(p)}{1!1!}(x-a)(y-b)+\frac{\frac{\partial^2 f}{\partial y^2}(p)}{0!\,2!}(y-b)^2$$
$$+\frac{\frac{\partial^3 f}{\partial x^3}(p)}{3!0!}(x-a)^3+\frac{\frac{\partial^3 f}{\partial y\partial x^2}(p)}{2!\,1!}(x-a)^2(y-b)+\frac{\frac{\partial^3 f}{\partial y^2\partial x}(p)}{1!\,2!}(x-a)(y-b)^2$$
$$+\frac{\frac{\partial^3 f}{\partial y^3}(p)}{0!3!}(y-b)^3+o((x-a)^4,(y-b)^4).$$

Procediendo inductivamente, es fácil ver que la parte cuártica del polinomio buscado es,

$$\triangleleft\ \frac{\frac{\partial^4 f}{\partial x^4}(p)}{4!0!}(x-a)^4+\frac{\frac{\partial^4 f}{\partial y\partial x^3}(p)}{3!1!}(x-a)^3(y-b)+\frac{\frac{\partial^4 f}{\partial y^2\partial x^2}(p)}{2!2!}(x-a)^2(y-b)^2$$
$$+\frac{\frac{\partial^4 f}{\partial y^3\partial x}(p)}{1!3!}(x-a)(y-b)^3+\frac{\frac{\partial^4 f}{\partial y^4}(p)}{0!4!}(y-b)^4\quad\triangleright$$

Podemos escribir la generalización de toda la discusión anterior en el siguiente resultado.

TEOREMA 4.5 (de Taylor) *Si $f(x,y)$ es un campo escalar de clase C^r, entonces alrededor del punto $p=(a,b)\in\Omega$ se tiene que la función se escribe como el polinomio de grado r en dos variables*

$$f(x,y)=f(p)+\sum_{k=1}^{r}\left(\sum_{i+j=k}\frac{\frac{\partial^k f}{\partial y^j\partial x^i}(p)}{i!j!}(x-a)^i(y-b)^j\right)$$
$$+o((x-a)^{r+1},(y-b)^{r+1})$$

donde la cantidad $o((x-a)^{r+1},(y-b)^{r+1})$ es un error de orden $r+1$ en los monomios $x-a$ y $y-b$.

DEFINICIÓN. A la expresión (sin error) del lado derecho de la ecuación en El teorema de Taylor se le llamará El **polinomio de Taylor** de grado r de la función f en el punto p, en las variables independientes.

EJEMPLO. Sea la función

$$f(x,y)=\operatorname{sen}(xy)$$

Desarrollar El polinomio de Taylor alrededor de $p=(0,0)$ hasta orden dos.

$\triangleleft$ Aquí, $a=0$, $b=0$, y al calcular las parciales (hasta las segundas) se tiene

$$\frac{\partial f}{\partial x}=y\cos(xy),\quad \frac{\partial^2 f}{\partial y\partial x}=\cos(xy)-xy\operatorname{sen}(xy)$$



$$\frac{\partial^2 f}{\partial x^2} = -y^2 \text{sen}(xy), \quad \frac{\partial f}{\partial y} = x\cos(xy), \quad \frac{\partial^2 f}{\partial y^2} = -x^2 \text{sen}(xy)$$

Por lo tanto, los coeficientes se calculan evaluando en el punto $(a, b) = (0, 0)$,

$$\frac{\partial f}{\partial x}(0,0) = 0, \ \frac{\partial^2 f}{\partial x^2}(0,0) = 0, \ \frac{\partial^2 f}{\partial y \partial x}(0,0) = 1, \ \frac{\partial f}{\partial y}(0,0) = 0, \ \frac{\partial^2 f}{\partial y^2}(0,0) = 0$$

y por la fórmula de Taylor del teorema 5.2 se tiene que,

$$f(x,y) = f(0,0) + \frac{\frac{\partial f}{\partial x}(0,0)}{1!0!}(x-0) + \frac{\frac{\partial f}{\partial y}(0,0)}{0!1!}(y-0) + \frac{\frac{\partial^2 f}{\partial x^2}(0,0)}{2!0!}(x-0)^2$$

$$+\frac{\frac{\partial^2 f}{\partial y \partial x}(0,0)}{1!1!}(x-0)(y-0) + \frac{\frac{\partial^2 f}{\partial y^2}(0,0)}{0!2!}(y-0)^2 + o((x-0)^3, (y-0)^3)$$

$$= xy + o(x^3, y^3)$$

Esto es, hasta orden dos la función dada se escribe

$$\text{sen}(xy) = xy + o(x^3, y^3)$$

Otra forma de realizar el desarrollo anterior, es observando que para una variable z real el desarrollo de la función seno alrededor de $z_0 = 0$ es

$$\text{sen} z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + o(8)$$

Ahora, defínase $z = xy$, y considérese $p = (0, 0)$. Entonces, $z(0, 0) = 0 = z_0$, y sustituyendo la variable z, hasta el orden dos, se obtiene

$$\text{sen}(xy) = (xy) - \frac{(xy)^3}{3!} + o(7) = xy + o(3)$$

que es la misma expresión obtenida. ▷

La forma de desarrollar las funciones en varias variables utilizando los desarrollos conocidos de Taylor para una variable, ayudan en general para omitir una buena cantidad de cálculos. No obstante hay que tener práctica para reconocer las sustituciones apropiadas y los puntos alrededor de los cuales se está desarrollando. Esto se ha mostrado en el ejemplo anterior y se muestra nuevamente en el siguiente.

EJEMPLO. Desarrollar la función

$$f(x, y) = \ln(1 + xy)$$

en polinomio de Taylor hasta orden tres alrededor del punto $p = (0,0)$.

◁ Calculamos las derivadas parciales hasta orden tres,

$$\frac{\partial f}{\partial x} = \frac{y}{1+xy}$$

$$\frac{\partial f}{\partial y} = \frac{x}{1+xy}$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{(1+xy) - yx}{(1+xy)^2} = \frac{1}{(1+xy)^2}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{-y^2}{(1+xy)^2}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{-x^2}{(1+xy)^2}$$

$$\frac{\partial^3 f}{\partial x^3} = \frac{2y^3}{(1+xy)^3}$$

$$\frac{\partial^3 f}{\partial y^3} = \frac{2x^3}{(1+xy)^3}$$

$$\frac{\partial^3 f}{\partial y^2 \partial x} = -2(1+xy)^{-3}x = \frac{-2x}{(1+xy)^3}$$

$$\frac{\partial^3 f}{\partial y \partial x^2} = \frac{-2y(1+xy)^2 + y^2 2(1+yx)x}{(1+xy)^4} = \frac{-2y}{(1+xy)^3}$$

De esta manera,

$$\frac{\partial f}{\partial x}(0,0) = 0, \ \frac{\partial f}{\partial y}(0,0) = 0$$

$$\frac{\partial^2 f}{\partial x^2}(0,0) = 0, \ \frac{\partial^2 f}{\partial y \partial x}(0,0) = 1, \ \frac{\partial^2 f}{\partial y^2}(0,0) = 0$$

$$\frac{\partial^2 f}{\partial x^3}(0,0) = 0, \ \frac{\partial^3 f}{\partial y \partial x}(0,0) = 0, \ \frac{\partial^3 f}{\partial y \partial x}(0,0) = 0, \ \frac{\partial^3 f}{\partial y^3}(0,0) = 0$$

Por lo tanto, alrededor de $p = (0,0)$ se tiene que la función $f(x,y) = \ln(1+xy)$ se escribe

$$\ln(1+xy) = \ln(1+0) + \frac{\partial f}{\partial x}(0,0)(x-0) + \frac{\partial f}{\partial y}(0,0)(y-0)$$

$$+\frac{\frac{\partial^2 f}{\partial x^2}(0,0)}{2!0!}(x-0)^2 + \frac{\frac{\partial^2 f}{\partial y \partial x}(0,0)}{1!1!}(x-0)(y-0) + \frac{\frac{\partial^2 f}{\partial y^2}(0,0)}{0!2!}(y-0)^2$$

$$+\frac{\frac{\partial^3 f}{\partial x^3}(0,0)}{3!0!}(x-0)^3+\frac{\frac{\partial^3 f}{\partial y\partial x^2}(0,0)}{2!1!}(x-0)(y-0)+\frac{\frac{\partial^3 f}{\partial y^2\partial x}(0,0)}{1!2!}(x-0)(y-0)^2$$

$$+\frac{\frac{\partial^3 f}{\partial y^3}(0,0)}{0!3!}(y-0)^3+o((x-0)^4,\,(y-0)^4)=xy+o(x^4,y^4)$$

Por otro lado, el desarrollo alrededor de $z_0=0$ de la función $\ln(1+z)$ está dado por la igualdad

$$\ln(1+z)=z-\frac{z^2}{2}+o(3)$$

Por eso, si ponemos $z=xy$ entonces el punto $(a,b)=(0,0)$ corresponde al punto $z_0=0$ y entonces

$$\ln(1+xy)=xy-\frac{(xy)^2}{2}+o(3)=xy-\frac{x^2y^2}{z}+o(5)$$

$$=xy+o(3)$$

igualdad que se ha obtenido ya, utilizando el teorema 4.5. ▷

Damos ahora un ejemplo del desarrollo de un polinomio alrededor de un punto.

EJEMPLO. Dar el polinomio de Taylor alrededor del punto $p=(1,2)$ de la función polinominal

$$f(x,y)=x^3-2y^3+3xy$$

◁ Calculamos todas las posibles derivadas parciales, obteniendo,

$$\frac{\partial f}{\partial x}=3x^2-3y,\ \frac{\partial f}{\partial y}=-6y^2+3x$$

$$\frac{\partial^2 f}{\partial x^2}=6x,\ \frac{\partial^2 f}{\partial y\partial x}=-3,\ \frac{\partial^2 f}{\partial y^2}=-12y$$

$$\frac{\partial^3 f}{\partial x^3}=6,\ \frac{\partial^3 f}{\partial y\partial x^2}=0,\ \frac{\partial^3 f}{\partial y^2\partial x}=0,\ \frac{\partial^3 f}{\partial y^3}=-12$$

De esta forma, al evaluar cada derivada obtenida en el punto p se tiene

$$\frac{\partial f}{\partial x}(1,2)=-3,\quad \frac{\partial f}{\partial y}(1,2)=-21$$

$$\frac{\partial^2 f}{\partial x^2}(1,2)=6,\quad \frac{\partial^2 f}{\partial y\partial x}(1,2)=-3,\quad \frac{\partial^2 f}{\partial y^2}(1,2)=-24$$



$$\frac{\partial^3 f}{\partial x^3}(1,2) = 6, \quad \frac{\partial^3 f}{\partial y \partial x^2}(1,2) = 0, \quad \frac{\partial^3 f}{\partial y^2 \partial x}(1,2) = 0, \quad \frac{\partial^3 f}{\partial y^3}(1,2) = -12$$

Por lo tanto, por el teorema 4.5, alrededor de $(1,2)$ la función f se escribe

$$\begin{aligned} f(x,y) = f(1,2) - \frac{-3(x-1)}{1!0!} - \frac{21(y-2)}{0!1!} \\ + \frac{6(x-1)^2}{2!0!} - \frac{3(x-1)(y-2)}{1!1!} - \frac{24(y-2)^2}{0!2!} \\ + \frac{6(x-1)^3}{3!0!} + \frac{0(x-1)^2(y-2)}{2!1!} + \frac{0(x-1)(y-2)^2}{1!2!} - \frac{12(y-2)^3}{0!3!} \\ = -9 - 3(x-1) - 21(y-2) + 3(x-1)^2 - 3(x-1)(y-2) - 12(y-2)^2 \\ +(x-1)^3 - 2(y-2)^3 \end{aligned}$$

debido a que las derivadas parciales de orden cuatro en adelante se anulan todas. ▷

4.8 Diferencial total de un campo escalar

Para el campo escalar $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ y el punto arbitrario $p = (x, y) \in D$, se tiene el **incremento total** de f en p dado por

$$\Delta f = f(x+h, y+k) - f(x, y)$$

donde $h = \Delta x$ y $k = \Delta y$ son incrementos arbitrarios en las variables dadas.

De la definición de diferenciabilidad de f en p, dada en la sección 4.4, se tiene que f es diferenciable en p si el incremento total Δf puede escribirse en la forma

$$\Delta f = \frac{\partial f}{\partial x}(p)\Delta x + \frac{\partial f}{\partial y}(p)\Delta y + o(||(\Delta x, \Delta y)||^2)$$

cuando $||(\Delta x, \Delta y)|| \to 0$. Podemos escribir lo anterior mediante la notación convencional

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

en virtud de que $o(||(\Delta x, \Delta y)||^2) \to 0$.

DEFINICIÓN. Para la función diferenciable $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ se define la **diferencial total** de f por la igualdad

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

EJEMPLO. Dada la función $f(x, y) = \ln(x^2 + y^2)$ definida en la región plana $D = \mathbb{R}^2 - \{(0, 0)\}$, calcule su diferencial total.

◁ En virtud de que

$$\frac{\partial f}{\partial x}(x, y) = \frac{2x}{x^2 + y^2}, \quad \frac{\partial f}{\partial y}(x, y) = \frac{2y}{x^2 + y^2}$$

se tiene

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = \frac{2x}{x^2 + y^2}dx + \frac{2y}{x^2 + y^2}dy = \frac{2xdx + 2ydy}{x^2 + y^2}$$

De esta manera, para calcular aproximadamente $\ln(0.091^2 + 0.99^2)$ tomamos $x = 0$ y $y = 1$ obteniendo que $\Delta x = 0.091$, $\Delta y = -0.01$, lo que nos indica que

$$df = \frac{2(0)(0.091) + 2(1)(-0.01)}{0^2 + 1^2} = -0.02$$

y consecuentemente,

$$\ln(0.091^2 + 0.99^2) \approx f(0, 1) + df = \ln(0^2 + 1^2) - 0.02$$

es decir,

$$\ln(0.091^2 + 0.99^2) \approx -0.02 \quad \triangleright$$

De manera análoga, para una función diferenciable de tres variables, se tiene una diferencial total,

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$$

EJEMPLO. Dada la función $T = e^{xyz}$, calcular su diferencial total.

◁ Primero calculamos las derivadas parciales

$$\frac{\partial T}{\partial x} = yze^{xyz}, \quad \frac{\partial T}{\partial y} = xze^{xyz}, \quad \frac{\partial T}{\partial z} = xye^{xyz}$$

luego construimos la diferencial total

$$dT = \frac{\partial T}{\partial x}dx + \frac{\partial T}{\partial y}dy + \frac{\partial T}{\partial z}dz = yze^{xyz}dx + xz\,e^{xyz}dy + xy\,e^{xyz}dz$$

$$= e^{xyz}(yzdx + xzdy + xydz) \quad \triangleright$$

Para una función de clase C^2 se define la **diferencial total de segundo orden** por $d^2 f$ y se calcula mediante

$$d^2 f = \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial y \partial x}dxdy + \frac{\partial^2 f}{\partial y^2}dy^2$$

Si escribimos la expresión anterior en notación operacional se tiene.

$$d^2 f = \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial y \partial x}dxdy + \frac{\partial^2 f}{\partial y^2}dy^2$$

$$= \left(\frac{\partial^2}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial y \partial x}dxdy + \frac{\partial^2}{\partial y^2}dy^2\right)(f)$$

$$= \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^2 (f)$$

donde la aplicación del operador se realiza mediante la convención

$$\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\right) = \frac{\partial^2}{\partial y \partial x} = \frac{\partial^2}{\partial x \partial y} = \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}\right)$$

EJEMPLO. Sea $u = \cos x \cos y$. Calcular $d^2 u$.

◁ Calculamos las derivadas parciales, obteniendo,

$$\frac{\partial u}{\partial x} = -\operatorname{sen} x \cos y, \quad \frac{\partial u}{\partial y} = -\cos x \operatorname{sen} y$$

$$\frac{\partial^2 u}{\partial x^2} = -\cos x \cos y, \quad \frac{\partial^2 u}{\partial y \partial x} = \operatorname{sen} x \operatorname{sen} y, \quad \frac{\partial^2 u}{\partial y^2} = -\cos x \cos y,$$

lo que indica que

$$du^2 = -\cos x \cos y dx^2 + 2 \operatorname{sen} x \operatorname{sen} y dxdy - \cos x \cos y dy^2 \quad \triangleright$$

De manera inductiva, se tiene para un campo escalar f de clase C^3 la tercera diferencial total $d^3 f = d(d^2 f)$, y viene dada por

$$d^3 f = \frac{\partial^3 f}{\partial x^3}dx^3 + 3\frac{\partial^3 f}{\partial y \partial x^2}dx^2 dy + 3\frac{\partial^3 f}{\partial y^2 \partial x}dxdy^2 + \frac{\partial^3 f}{\partial y^3}dy^3$$

que en notación convencional operacional podemos escribir

$$d^3 f = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^3 (f)$$

En forma general, en notación convencional, se tiene la $k-$**ésima diferencial total** del campo escalar de clase C^k dada por

$$d^k f = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^k (f)$$

que se desarrolla de acuerdo a la fórmula del binomio de Newton.

EJEMPLO. Para el campo escalar $u = y\ln x$ calcular d^4u.

$\lhd$ De acuerdo al Triángulo de Pascal, los coeficientes del binomio de Newton de grado 4 son 1, 4, 6, 4, 1, lo cual nos indica que

$$d^4u = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^4 (u)$$

$$= \left(\frac{\partial^4}{\partial x^4}dx^4 + 4\frac{\partial^4}{\partial y\partial x^3}dx^3dy + 6\frac{\partial^4}{\partial y^2\partial x^2}dx^2dy^2 + 4\frac{\partial^4}{\partial y^3\partial x}dxdy^3 + \frac{\partial^4}{\partial y^4}dy^4\right)(u)$$

$$= \frac{\partial^4 u}{\partial x^4}dx^4 + 4\frac{\partial^4 u}{\partial y\partial x^3}dx^3dy + 6\frac{\partial^4 u}{\partial y^2\partial x^2}dx^2dy^2 + 4\frac{\partial^4 u}{\partial y^3\partial x}dxdy^3 + \frac{\partial^4 u}{\partial y^4}dy^4$$

De esta manera, al calcular las derivadas parciales

$$\frac{\partial u}{\partial x} = \frac{y}{x},\ \frac{\partial u}{\partial y} = \ln x$$

$$\frac{\partial^2 u}{\partial x^2} = -\frac{y}{x^2},\ \frac{\partial^2 u}{\partial y\partial x} = \frac{1}{x},\ \frac{\partial^2 u}{\partial y^2} = 0$$

$$\frac{\partial^3 u}{\partial x^3} = \frac{2y}{x^3},\ \frac{\partial^3 u}{\partial y\partial x^2} = -\frac{1}{x^2},\ \frac{\partial^3 u}{\partial y^2\partial x} = 0,\ \frac{\partial^3 u}{\partial y^3} = 0$$

$$\frac{\partial^4 u}{\partial x^4} = -\frac{6y}{x^4},\ \frac{\partial^4 u}{\partial y\partial x^3} = \frac{2}{x^3},\ \frac{\partial^4 u}{\partial y^2\partial x^2} = \frac{\partial^4 u}{\partial y^3\partial x} = \frac{\partial^4 u}{\partial y^4} \equiv 0$$

y sustituirlas se obtiene

$$d^4u = -\frac{6y}{x^4}dx^4 + 4\left(\frac{2}{x^3}\right)dx^3dy = -\frac{6y}{x^4}dx^4 + \frac{8}{x^3}dx^3dy \quad \rhd$$

Si consideramos el **elemento vectorial diferencial** en $\mathbb{R}^3$, en la base canónica $\{e_1, e_2, e_3\}$ dado por el vector simbólico

$$d\vec{r} = (dx, dy, dz)$$

entonces la diferencial total de un campo escalar f está dada por

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z} = \left\langle \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right), (dx, dy, dz)\right\rangle$$

$$= <\operatorname{grad} f,\, d\vec{r}>$$



Ejercicios

1. ¿Cuáles de los siguientes subconjuntos son abiertos? ¿Cuáles son cerrados? ¿Cuáles son conexos? Haga un dibujo para cada conjunto dado.

a. $\Omega = \{(x,y) |\ |x| \geq 1\}$
b. $\Omega = \{(x,y) |\ x + y < 1\}$
c. $\Omega = \{(x,y) |\ |x| + |y| \geq 1\}$
d. $\Omega = \{(x,y) |\ |\operatorname{sen} x| > \frac{\sqrt{2}}{2}\}$
e. $\Omega = \{(x,y,z) |\ x + y + 2z < 2\}$
f. $\Omega = \{(x,y,z) |\ x^2 - y^2 - z^2 \leq 1\}$
g. $\Omega = \{(x,y,z) |\ ||(x,y,z)|| < 1 \quad \text{ó} \quad 4 \leq ||(x,y,z)||\}$
h. $\Omega = \{(x,y,z) |\ x^2 + y^2 < z < \sqrt{4 - x^2 - y^2}\}$

2. Calcular los dominios de los siguientes campos escalares y dar su característica topológica (abierto, cerrado, conexo).

a. $f_1(x,y) = \sqrt{x^2 + y^2 - 4}$
b. $f_2(x,y) = \frac{4}{\sqrt{4 - x^2 y^2}}$
c. $f_3(x,y) = \arccos(2x - y)$
d. $f_4(x,y) = \ln(y^2 - x)$
e. $f_5(x,y) = y + \sqrt{x}$
f. $g_1(x,y,z) = \sqrt{16 - x^2 - y^2 - z^2}$
g. $g_2(x,y,z) = \operatorname{arcsen}\left(\frac{3x}{\sqrt{y^2 + z^2}}\right)$
h. $g_3(x,y,z) = \frac{1}{\ln(9 - x^2 - y^2 - z^2)}$
i. $g_4(x,y,z) = \sqrt{x + 2y - 3z}$

3. Dibujar las curvas de nivel de las siguientes funciones definidas en dominios del plano. Hallar tales dominios.

a. $f_1(x,y) = 4x - y$
b. $f_2(x,y) = \frac{2y}{x}$
c. $f_3(x,y) = \ln\left(\frac{1-x}{1-y}\right)$
d. $f_4(x,y) = e^{x+y}$

4. Hallar las superficies de nivel de las siguientes funciones definidas en dominios de $\mathbb{R}^3$.

a. $f_1(x,y,z) = 6x - 3y + 2z$



b. $f_2(x,y,z) = 3x^2 + 3y^2 + z$
c. $f_3(x,y,z) = \sqrt{x^2 + y^2 + z^2}$
d. $f_4(x,y,z) = x^2 - y^2 - z^2$

5. Calcular las derivadas parciales $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial}{\partial z}$ de cada una de las funciones indicadas.

a. $f_1(x,y,z) = x^3y^4 + xyz^2 - 3z^3y^3$
b. $f_2(x,y,z) = \operatorname{sen}(xyz) + x\operatorname{sen} y + y\operatorname{sen} z$
c. $f_3(x,y,z) = x\operatorname{sen}(y-z) + z\arctan(xy)$
d. $f_4(x,y,z) = \sqrt{x^2+y^2}e^{yz} + \sqrt{x^2+z^2}\ln(xy)$
e. $f_5(x,y,z) = \frac{x-y+z}{x^2+y^2z^2}$

6. a. Mostrar que la función $z = y^{y/x}\operatorname{sen}\left(\frac{y}{x}\right)$ satisface la ecuación

$$x^2\frac{\partial z}{\partial x} + xy\frac{\partial z}{\partial y} = yz$$

b. Probar que la función $z = \frac{x^2}{2y} + \frac{x}{2} + \frac{1}{x} - \frac{1}{y}$ satisface la ecuación

$$x^2\frac{\partial z}{\partial x} + y^2\frac{\partial z}{\partial y} = \frac{x^3}{y}$$

7. Para los incisos del ejercicio **5.** correspondientes, calcular el gradiente de la función correspondiente en el punto dado.

a. $\nabla f_1(x,y,z)$ en $p = (1,1,1)$
b. $\nabla f_2(x,y,z)$ en $p = \left(1, \frac{\pi}{2}, \frac{\pi}{4}\right)$
c. $\nabla f_3(x,y,z)$ en $p = \left(0, \frac{\pi}{2}, \frac{\pi}{4}\right)$
d. $\nabla f_4(x,y,z)$ en $p = (-1,0,1)$
e. $\nabla f_5(x,y,z)$ en $p = (1,2,3)$

8. a. Sean

$$\begin{cases} x = t + 2s \\ y = s^2 - t \\ z = t^2 + s^2 \end{cases}$$

y la función $u = x^2 + 2xy + y^2 + z^2$. Calcular $\frac{\partial u}{\partial s}$ y $\frac{\partial u}{\partial t}$ mediante la regla de la cadena.

b. Dados

$$x = \operatorname{sen}(t+s), \quad y = \cos(t+s)$$

y la función $u = \frac{x+xy}{x^2+y^2}$, calcular $\frac{\partial u}{\partial t}$, $\frac{\partial u}{\partial s}$ mediante la regla de la cadena.



c. Dado $u = (x^2 + y^2 + z^2)^{3/2}$, calcular

$$\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}$$

d. Sea la función $u = f(r - s,\, s - t,\, t - r)$, demostrar que

$$\frac{\partial u}{\partial s} + \frac{\partial u}{\partial r} + \frac{\partial u}{\partial t} = 0$$

e. Sea $u = f\left(\frac{x^2 - y^2}{x^2 + y^2}\right)$, demostrar que

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = 0$$

f. Dada la función $g(x, y) = f(x + y,\, x - y)$, demostrar la igualdad

$$\frac{\partial g}{\partial x}\frac{\partial g}{\partial y} = \left(\frac{\partial f}{\partial u}\right)^2 - \left(\frac{\partial f}{\partial v}\right)^2$$

donde $u = x + y$, $v = x - y$.

g. Dado el ángulo constante φ y las relaciones

$$\begin{cases} x = u\cos\varphi - v\,\mathrm{sen}\,\varphi \\ y = u\,\mathrm{sen}\,\varphi + v\cos\varphi \end{cases}$$

demostrar que la función compuesta $f(x, y) = g(u, v)$ satisface la relación

$$\left(\frac{\partial g}{\partial u}\right)^2 + \left(\frac{\partial g}{\partial v}\right)^2 = \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2$$

h. Calcular el gradiente ∇f_p para las siguientes funciones, si

$$r = \sqrt{x^2 + y^2 + z^2} \text{ y } f(x, y, z) = h(r)$$

donde,

i) $h(r) = r^2$
ii) $h(r) = \ln r$
iii) $h(r) = e^{-r^2}$

9. Calcular la derivada direccional $D_\xi f_p$ para las funciones, puntos y vectores dados.



a. $f(x,y) = xy^2, \quad p = (1,1), \quad \xi = (-1,3)$

b. $f(x,y) = \arctan(xy), \quad p = (-1,1), \quad \xi = (1,-2)$

c. $f(x,y) = e^{-(x^2+y^2)}, \quad p = (0,0), \quad \xi = (1,1)$

d. $f(x,y,z) = y \arctan(xz), \quad p = (1,1,1), \quad \xi = (-1,0,2)$

e. $f(x,y,z) = xy + yz + zx, \quad p = (-1,1,-1), \quad \xi = (1,-1,2)$

10. a. Dada la función $f(x,y,z) = x^2 + xy + z^2$ dar la dirección ξ en la que crece más rápidamente f en el punto $p = (-1,1,1)$. Calcular $D_\xi\, f_p$ para estos elementos.

b. Sea la función

$$f(x,y) = \frac{x}{\sqrt{x^2+y^2}}$$

En qué dirección crece más rápidamente f en el punto $p = (1,0)$?

11. a. Calcular $\frac{\partial^2 f}{\partial y \partial x}$ y $\frac{\partial^2 f}{\partial x \partial y}$ para $f(x,y) = \ln\tan(y+x)$

b. Calcular $\frac{\partial^2 f}{\partial y \partial x}$ y $\frac{\partial^2 f}{\partial x \partial y}$ para $f(x,y) = \arctan\left(\frac{x+y}{1+xy}\right)$

c. Calcular $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2}$ y $\frac{\partial^2 f}{\partial y \partial x}$ para $f(x,y) = y \operatorname{sen} xy + x \cos xy$

d. Calcular $\frac{\partial^3 u}{\partial y \partial x^2}, \frac{\partial^3 u}{\partial y^2 \partial x}$ y $\frac{\partial^3 u}{\partial x^2 \partial y}$ para $u = \operatorname{sen}(2x + y^2)$

e. Dada la función $u = x^3y^2$ demostrar la validez de la igualdad

$$\frac{\partial^5 u}{\partial y^2 \partial x^3} = \frac{\partial^5 u}{\partial x^3 \partial y^2}$$

f. Sea la función $u = \frac{1}{\sqrt{t}} e^{-\frac{x^2}{4t^2}}$. Probar que satisface la ecuación

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

12. Sea $f(x,y) = g(r,\varphi)$ donde

$$\begin{cases} x = r\cos\varphi \\ y = r\operatorname{sen}\varphi \end{cases}$$

demostrar que se cumple la igualdad

$$\frac{\partial^2 g}{\partial r^2} + \frac{1}{r}\frac{\partial g}{\partial r} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$



13. Sea $f(x,y)$ una función tal que para cualquier número real t se cumple que en todo el dominio de f es justa la igualdad $f(tx,ty) = t^n f(x,y)$, donde n es un número entero.

a. Demostrar que se cumple la igualdad

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = nf(x,y)$$

b. Mostrar además que también se cumple que

$$x^2\frac{\partial^2 f}{\partial x^2} + 2xy\frac{\partial^2 f}{\partial x \partial y} + y^2\frac{\partial^2 f}{\partial y^2} = n(n-1)f(x,y)$$

14. Demostrar que la función $u = e^{3x+4y}\cos(5z)$ satisface La **ecuación diferencial de Laplace**,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

15. Desarrollar hasta orden tres en fórmula de Taylor las funciones indicadas en los puntos indicados.

a. $f(x,y) = e^{2x}\operatorname{sen} 2y$ en el punto $p = (1, \pi/2)$.
b. $f(x,y) = \cos(x^2 - y)$ en el punto $p = (0, -\pi)$.
c. $f(x,y) = \ln(1 + x + y)$ en el punto $p = (1,1)$.
d. $f(x,y) = x^2 + xy + y^3$ en el punto $p = (0,0)$
e. $f(x,y) = x^2 + xy + y^2$ en el punto $p = (-1,1)$

16. Dado el polinomio

$$f(x,y) = y^3 + 4xy^5 + 4x^2y^4$$

¿Cuál es el término de orden cuatro de su expresión de Taylor en el punto $p = (0,0)$?

17. Calcular df, df^2 y $d^3 f$ para las funciones **a.-e.** del ejercicio **15.**



Capítulo 5

Campos vectoriales en $\mathbb{R}^3$

5.1 Funciones del tipo $\mathbb{R}^n \to \mathbb{R}^m$

En este capítulo hacemos un estudio de los aspectos básicos de las funciones vectoriales con argumento vectorial del tipo $\mathbb{R}^n \to \mathbb{R}^m$, llamadas también **campos vectoriales** definidos en $\mathbb{R}^n$ con valores vectoriales en $\mathbb{R}^m$. De hecho, los dominios de tales funciones en general serán regiones contenidas en el espacio de dimensión n.

Dada la región (conjunto abierto conexo) $\Omega \subset \mathbb{R}^n$, una función de tipo

$$f : \Omega \to \mathbb{R}^m$$

a cada punto $p \in \Omega$ le asocia el punto $q = f(p)$ en $\mathbb{R}^m$.

En los correspondientes sistemas de coordenadas, si $p = (x_1, \cdots, x_n)$, $q = (y_1, \cdots, y_m)$ entonces tal función se escribirá

$$(y_1, y_2, \cdots, y_m) = f(x_1, \cdots, x_n)$$

Como las coordenadas y_i's dependen de p, entonces cada una de ellas es una función real dependiente de p,

$$\begin{cases} y_1 = f_1(p) = f_1(x_1, \cdots, x_n) \\ y_2 = f_2(p) = f_2(x_1, \cdots, x_n) \\ \qquad \vdots \\ y_m = f_m(p) = f_m(x_1, \cdots, x_n) \end{cases}$$

siendo cada una de las fi, funciones reales de variable vectorial definidas en Ω, para $i = 1, 2, \cdots, m$.

$$f_i : \Omega \to \mathbb{R}$$



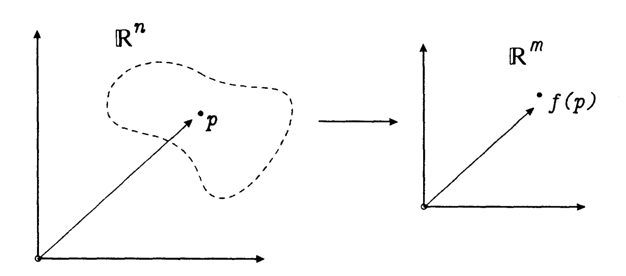


Figura 5.1: Función de $\mathbb{R}^n \to \mathbb{R}^m$.

y son llamadas las **funciones coordenadas** de la función f (véase la figura 5.1).

Esto es, la función f se escribe en coordenadas, como

$$f(x_1, \cdots, x_n) = (f_1(x_1, \cdots, x_n), f_2(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n))$$

EJEMPLO. Sea la función f cuya regla de correspondencia está dada por

$$f(x, y) = \left(xy, \ln xy^2, \frac{y}{1+x^2}\right),$$

Entonces f está definida en una región Ω de $\mathbb{R}^2$ y toma valores en $\mathbb{R}^3$.

$\lhd$ Procedamos a calcular el dominio de la función, analizando cada función coordenada.

Ciertamente la coordenada primera xy y la tercera $\frac{y}{1+x^2}$ no ofrecen alguna restricción. La segunda función coordenada impone la condición de que el argumento xy^2 deberá ser positivo.

Como $xy^2 > 0$, necesariamente cuando $x > 0$, $y \neq 0$, el dominio es entonces

$$\Omega = \{(x, y) | \, x > 0, \, y \neq 0\}$$

y la función f está bien definida en Ω. Las funciones coordenadas se escriben explícitamente,

$$f_1(x, y) = xy, \quad f_2(x, y) = \ln xy^2, \quad f_3(x, y) = \frac{y}{1+x^2} \quad \rhd$$

La figura 5.2 ilustra el dominio de la función. $\rhd$

EJEMPLO. Sea la función vectorial de argumento real

$$\gamma(t) = (t^2 + 1, \sqrt{1-t}, \operatorname{sen} t)$$



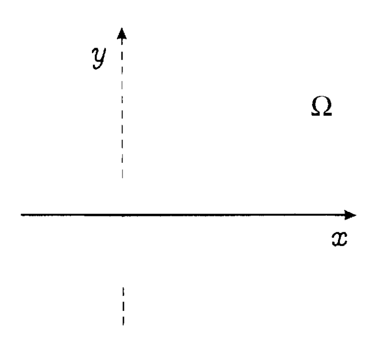


Figura 5.2: Dominio de $(x, y) \to \left(xy, \ln xy^2, \frac{y}{1+x^2}\right)$.

Claramente, es una función real con valores en $\mathbb{R}^3$, esto es, tiene la forma $\Omega \subset \mathbb{R} \to \mathbb{R}^3$.

$\lhd$ Se condiciona la variable t a que $1 - t \geq 0$, lo que nos dice que $1 \geq t$, de donde el dominio de γ es entonces,

$$\Omega = \{t \in \mathbb{R} \mid t \leq 1\} = (-\infty, 1]$$

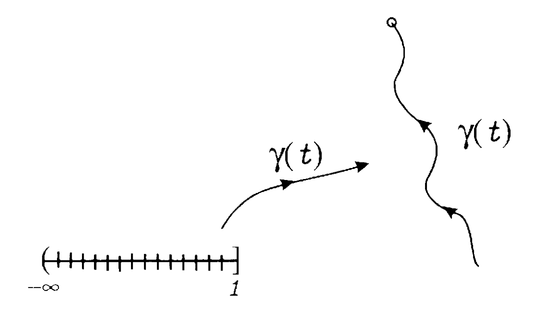


Figura 5.3: Curva espacial.

Las funciones coordenadas de γ vienen dadas por

$$f_1(t) = t^2 + 1, \quad f_2(t) = \sqrt{1-t}, \quad f_3(t) = \operatorname{sen} t$$

La figura 5.3 ilustra la imagen de γ en el espacio de dimensión tres. $\rhd$

EJEMPLO. Sea la función dada por

$$f(x, y, z) = (x^2 + y^2, \ln xz + y)$$



Esta función está definida en una región espacial y toma valores en $\mathbb{R}^2$, esto es, f es de la forma

$$\Omega \subset \mathbb{R}^3 \to \mathbb{R}^2.$$

◁ Para calcular su dominio imponemos la condición $xz > 0$, lo que equivale a

$$\begin{cases} x > 0 \quad \text{y} \quad z > 0 \\ \qquad \text{ó} \\ x < 0 \quad \text{y} \quad z < 0 \end{cases}$$

lo que implica que el dominio es la región espacial

$$\Omega = \{(x, y, z) \mid x > 0 \quad \text{y} \quad z > 0, \quad \text{ó} \quad x < 0 \quad \text{y} \quad z < 0, \quad y \in \mathbb{R}\}$$

La figura 5.4 ilustra el dominio de la función, que es la unión de cuatro octantes en $\mathbb{R}^3$.

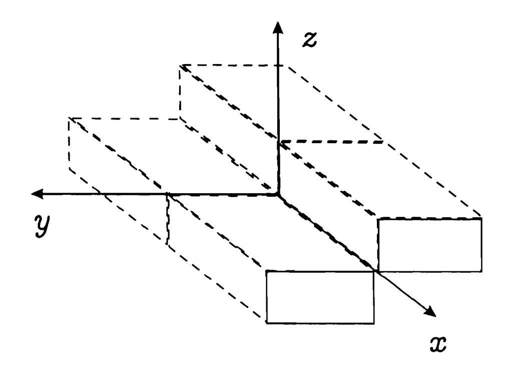


Figura 5.4: Dominio del campo $(x^2 + y^2, \ln xz + y)$.

Las funciones coordenadas del **campo vectorial** f son, en este caso

$$f_1(x, y, z) = x^2 + y^2, \quad f_2(x, y, z) = \ln xz + y \quad \triangleright$$

EJEMPLO. El campo escalar $f : \mathbb{R}^3 \to \mathbb{R}$ dado por la regla de correspondencia

$$f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$$

está definido en $\mathbb{R}^3$ $(n = 3, m = 1)$, y tiene sólo una función coordenada.

Le damos una estructura aritmética al conjunto de funciones del tipo $\Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ mediante las propiedades vectoriales del codominio $\mathbb{R}^m$.



Si $f, g : \Omega \to \mathbb{R}^m$ son dos funciones definidas en el mismo dominio en $\mathbb{R}^n$, se puede definir la suma de ellas

$$f + g : \Omega \to \mathbb{R}^m$$

como la función que en coordenadas del dominio y codominio se calcula por

$$(f + g)(x_1, \cdots, x_n) = f(x_1, \cdots x_n) + g(x_1, \cdots, x_n)$$

$$= (f_1(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n))$$
$$(g_1(x_1, \cdots, x_n), \cdots, g_m(x_1, \cdots, x_n))$$

$$= (f_1(x_1, \cdots, x_n) + g_1(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots x_n) + g_m(x_1, \cdots, x_n))$$

Si λ es un escalar, se define la función nueva

$$\lambda f : \Omega \to \mathbb{R}^m$$

como aquella que en coordenadas se escribe

$$(\lambda f)(x_1, \cdots, x_n) = \lambda f(x_1, \cdots, x_n) = \lambda(f_1(x_1, \cdots x_n), \cdots, f_m(x_1, \cdots, x_n))$$
$$= (\lambda f_1(x_1, \cdots, x_n), \cdots, \lambda f_m(x_1, \cdots, x_m))$$

También definimos una categoría de las funciones vectoriales mediante la composición ordinaria de funciones.

Sean las funciones $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, $g : U \subset \mathbb{R}^m \to \mathbb{R}^k$ de tal forma que la imagen de Ω bajo f esté contenida en U, es decir, $f(\Omega) \subset U$. Se define la composición de la función f con g

$$g \circ f : \Omega \to \mathbb{R}^k$$

a través de la igualdad,

$$(g \circ f)(p) = g(f(p))$$

véase la figura 5.5.

A continuación damos ejemplos de la composición de funciones vectoriales de argumento vectorial.

EJEMPLO. Sean las funciones vectoriales dadas por

$$f(x, y, z) = (x^2, y^2 + z^2)$$



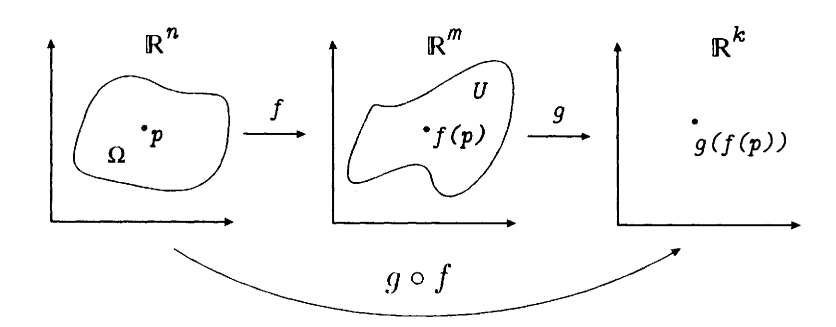


Figura 5.5: Composición de funciones.

$$g(u,v) = (\ln(u+v), e^{uv}, uv)$$

◁ **a.** Calculemos $g \circ f$ en un punto arbitrario (x, y, z) del dominio de esta composición.

$$(g \circ f)(x,y,z) = g(f(x,y,z)) = g(x^2, y^2+z^2)$$
$$= (\ln(x^2+y^2+z^2), e^{x^2(y^2+z^2)}, x^2(y^2+z^2))$$

de donde, la composición es del tipo

$$g \circ f : \Omega_1 \subset \mathbb{R}^3 \to \mathbb{R}^3$$

b. Calculamos $f \circ g$ en un punto arbitrario (u, w) del dominio correspondiente.

$$(f \circ g)(u,v) = f(g(u,v)) = f(\ln(u+v), e^{uv}, uv) = (\ln^2(u+v), (e^{uv})^2 + (uv)^2)$$
$$= (\ln^2(u+v), e^{2uv} + u^2v^2)$$

composición que tiene el tipo

$$f \circ g : \Omega_2 \subset \mathbb{R}^2 \to \mathbb{R}^2 \quad \triangleright$$

EJEMPLO. Dadas las funciones vectoriales

$$f(x,y) = \left(\frac{y}{1+x^2}, xy + y^2\right)$$

$$g(u,v,w) = (uv - w, u + v + w)$$

Calcular $g \circ f$, $f \circ g$.



$\triangleleft$ **a.** Ciertamente $g \circ f$ no tiene sentido pues la imagen de f es un subconjunto del plano $\mathbb{R}^2$, mientras que el dominio de g es la región espacial $\mathbb{R}^3$.

b. Para el punto (u, v, w) se cumple que

$$(f \circ g)(u, v, w) = f(g(u, v, w)) = f(uv - w, u + v + w)$$

$$= \left(\frac{u + v + w}{1 + (uv - w)^2}, (uv - w)(u + v + w) + (u + v + w)^2 \right) \quad \triangleright$$

Para optimizar la discusión sobre la continuidad de las funciones vectoriales de argumento vectorial, apelamos a la siguiente definición que garantiza la continuidad de los campos vectoriales en los subsecuentes ejemplos.

DEFINICIÓN. Una función $f : \Omega \to \mathbb{R}^n \to \mathbb{R}^m$ es continua en Ω, si y sólo sí, cada una de sus funciones coordenadas $f_i : \Omega \to \mathbb{R}$ es continua.

Transformaciones Lineales

De entre todas las funciones $\mathbb{R}^n \to \mathbb{R}^m$ se destacan las llamadas **tranformaciones lineales**. Estas ayudarán a definir la diferenciabilidad de una función vectorial de variable vectorial.

Iniciamos el tratado básico con las transformaciones lineales más simples $\mathbb{R}^3 \to \mathbb{R}^3$, obteniendo la mayor información para este tipo de transformaciones y argumentamos inductivamente la generalización de estos resultados para transformaciones lineales más generales del tipo $\mathbb{R}^n \to \mathbb{R}^m$.

Sea la función $L : \mathbb{R}^3 \to \mathbb{R}^3$ dada por

$$L(x, y, z) = (a_{11}x + a_{12}y + a_{13}z,\ a_{21}x + a_{22} + a_{23}z,\ a_{31}x + a_{32}y + a_{33}z)$$

donde cada a_{ij} es un número real.

$\triangleleft$ Es claro que si u, v, w son las coordenadas, en el codominio de L, entonces cada función se escribirá,

$$u = a_{11}x + a_{12}y + a_{13}z$$

$$v = a_{21}x + a_{22}y + a_{23}z$$

$$w = a_{31}x + a_{32}y + a_{33}z$$

que matricialmente se escribe,

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$



En otras palabras, si definimos a los vectores variable dependiente e independiente como los vectores columna,

$$p = \begin{pmatrix} x \\ y \\ z \end{pmatrix},\ A = (a_{ij}),\ q = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

la relación anterior se escribe mediante la igualdad

$$q = Ap = L(p)$$

En adelante, abusando de la notación, cada vez que en una transformación lineal apliquemos a un vector, para el cálculo tomaremos a tal vector como un vector columna, escribiendo finalmente el resultado del cálculo como un vector renglón elemento del codominio.

Con esta asociación y con la conversón de notación se tiene,

a. Si $p_1, p_2 \in \mathbb{R}^3$ (considerados como vectores), entonces

$$L(p_1 + p_2) = A(p_1 + p_2) = Ap_1 + Ap_2 = L(p_1) + L(p_2)$$

es decir, se cumple que

$$L(p_1 + p_2) = L(p_1) + L(p_2)$$

b. Si $\lambda \in p$ entonces $L(\lambda p) = A(\lambda p) = \lambda(Ap) = \lambda L(p)$, es decir, se satisface la igualdad,

$$L(\lambda p) = \lambda L(p)$$

DEFINICIÓN. La función L se dice que es una **transformación lineal** por satisfacer las propiedades **a.** y **b.**

Hemos observado que una matriz $A = A_{3\times 3}$ tiene asociada una transformación lineal $L : \mathbb{R}^3 \to \mathbb{R}^3$ que satisface las propiedades **a.** y **b.** de la definición.

Recíprocamente, sea dada la transformación $L : \mathbb{R}^3 \to \mathbb{R}^3$ tal que satisface **a.** y **b.** (es decir, es **lineal**), y sean los vectores $\{e_1, e_2, e_3\}$ básicos en $\mathbb{R}^3$ tal que todo punto (x, y, z) se escriba

$$(x, y, z) = xe_1 + ye_2 + ze_3$$

Definamos a los vectores columna (escritos convencionalmente de esta manera),

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} = L(e_1), \quad \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} = L(e_2), \quad \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} = L(e_3)$$



Afirmamos que la matriz construida por los vectores columna así obtenidos

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

está asociada a la tranformación lineal L y la define cuando se evaluan los puntos del dominio

$\triangleleft$ En efecto, al tomar un punto (x, y, z) se tiene

$$L(x, y, z) = L(xe_1 + ye_2 + ze_3) = L(xe_1) + L(ye_2) + L(ze_3)$$

$$= xL(e_1) + yL(e_2) + zL(e_3)$$

$$= x \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + y \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + z \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \triangleright$$

Podemos generalizar esto para dimensiones arbitrarias m, n haciendo una reconstrucción similar a la utilizada para este caso especial. Una **tranformación lineal** $L : \mathbb{R}^n \to \mathbb{R}^m$ es aquel campo vectorial que satisface las condiciones **a.** y **b.** de la definición dada. De esta manera se tendrá el siguiente teorema general.

TEOREMA 5.1 *Sea $L : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal, entonces existe una única matriz real de $m \times n$, $A = (a_{ij})$ asociada a L tal que para todo $x = (x_1, \cdots, x_n)$ se tiene que (utilizando la convención para el cálculo).*

$$L(x) = Ax = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

La matriz A es tal que si $\{e_1, e_2, \cdots e_n\}$ es la base en $\mathbb{R}^n$, entonces, L aplicado al vector e_j es el vector columna,

$$L(e_j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

esto es, el $j-$ésimo vector columna de la matriz asociada A.



◁ La unicidad del teorema se deja como un ejercicio para el lector. ▷

Hasta aquí, el lector ya habrá observado que una transformación lineal en general, tendrá sólo en sus funciones coordenadas a expresiones lineales en las variables, sin términos independientes.

En los siguientes ejemplos el lector deberá aprender a reconocer por simple inspección una transformación lineal, y a calcular su matriz asociada.

EJEMPLO. Sea la transformación $L : \mathbb{R}^3 \to \mathbb{R}^3$ dada por,

$$L(x, y, z) = (y + z,\ x + z,\ x + y)$$

◁ Al considerar la matriz A dada por

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

y a (x, y, z) un vector en $\mathbb{R}^3$, se tiene,

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y + z \\ x + z \\ x + y \end{pmatrix}$$

y, por lo tanto, A está asociada a L, lo que dice que L es una transformación lineal.

Más aún, si se toma la base canónica

$$\{e_1 = (1, 0, 0),\ e_2 = (0, 1, 0),\ e_3 = (0, 0, 0)\}$$

entonces, al aplicar A en cada uno de ellos se obtiene

$$A(e_1) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$A(e_2) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

$$A(e_3) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

lo que nos da los vectores columnas de la matriz A. ▷


EJEMPLO. Sea la transformación $L : \mathbb{R}^2 \to \mathbb{R}^3$ definida por

$$L(x, y) = (2x - y,\ 3x + 4y,\ 2x)$$

◁ Al considerar la matriz de 3×2 dada por

$$A = \begin{pmatrix} 2 & -1 \\ 3 & 4 \\ 2 & 0 \end{pmatrix}$$

se tiene que

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 3 & 4 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - y \\ 3x + 4y \\ 2x \end{pmatrix}$$

lo que indica que A está asociada a L.

Si se toman los vectores básicos $e_1 = (1, 0)$, $e_2 = (0, 1)$, entonces tendremos los vectores columna de la matriz A mediante los cálculos.

$$L(e_1) = \begin{pmatrix} 2 & -1 \\ 3 & 4 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}$$

$$L(e_2) = \begin{pmatrix} 2 & -1 \\ 3 & 4 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ 0 \end{pmatrix} \quad ▷$$

EJEMPLO. Sea la transformación lineal

$$L(x, y, z) = (-x - y + z,\ 4z - 2x + 2y)$$

◁ Al considerar la matriz real

$$A = \begin{pmatrix} -1 & -1 & 1 \\ -2 & 2 & 4 \end{pmatrix}_{2\times 3}$$

es fácil ver que A es la matriz asociada a L ▷

EJEMPLO. Sea la transformación lineal $L : \mathbb{R}^3 \to \mathbb{R}$ dada por

$$L(x, y, z) = 6x - 3y + 2z$$

◁ Si se considera el vector renglón,

$$A = \begin{pmatrix} 6 & -3 & 2 \end{pmatrix}_{1\times 3}$$

para el vector (x, y, z), se obtiene,

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (6 \quad -3 \quad 2)_{1\times 3}\begin{pmatrix} x \\ y \\ z \end{pmatrix}_{3\times 1} = 6x - 3y + 2z$$

lo que dice que A es la matriz asociada a la transformación lineal L. ▷

Damos ahora una interpretación geométrica de una transformación lineal para el caso $L : \mathbb{R}^2 \to \mathbb{R}^3$.

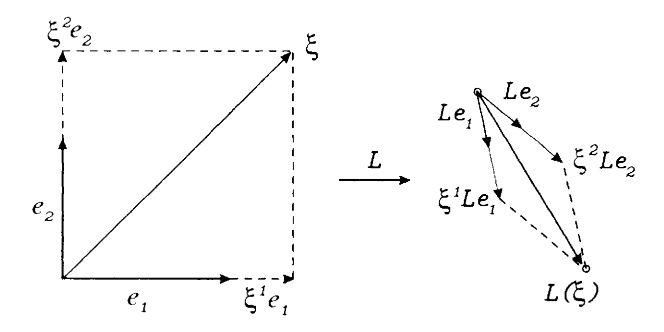


Figura 5.6: Geometría de una transformación lineal.

Si consideramos un vector en el plano $\mathbb{R}^2$, este se puede escribir como una **combinación lineal** de los vectores básicos,

$$\xi = \xi^1 e_1 + \xi^2 e_2$$

De esta manera, su imagen bajo L se escribiría como

$$L(\xi) = L(\xi^1 e_1 + \xi^2 e_2) = L(\xi^1 e_1) + L(\xi^2 e_2) = \xi^1 L(e_1) + \xi^2 L(e_2)$$

lo que nos dice que la imagen de ξ es una combinación lineal de los vectores imágenes de los básicos, en $\mathbb{R}^3$ (véase la figura 5.6).

De esta manera, la imagen del cuadrado en el plano $\mathbb{R}^2$ generado por los vectores e_1, e_2 bajo la transformación L es el paralelogramo en $\mathbb{R}^3$ tendido sobre los vectores $L(e_1)$ y $L(e_2)$, como lo muestra la figura 5.6.

En general, si $L : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal y $\xi \in \mathbb{R}^n$ es un vector posicionado en el origen, entonces $L(\xi)$ es también un vector en $\mathbb{R}^m$ posicionado en el origen.

Obtenemos ahora una matriz asociada a la composición de transformaciones lineales, en términos de las matrices asociadas a cada una de las

transformaciones que intervienen en la composición. Lo hacemos apenas para el caso especial de las transformaciones lineales del tipo $\mathbb{R}^2 \to \mathbb{R}^2$, generalizando el resultado de manera natural.

Consideremos dos transformaciones lineales del plano en el plano, dadas por

$$L_1(x,y) = (ax + by, cx + dy)$$

$$L_2(u,v) = (\alpha u + \beta v,\ \gamma u + \delta v)$$

donde $a, b, c, d, \alpha, \beta, \gamma, \delta$ son números reales.

$\triangleleft$ Al realizar su composición tenemos para un vector $(x,y) \in \mathbb{R}^2$,

$$(L_2 \circ L_1)(x,y) = L_2(L_1(x,y)) = L_2(ax + by, cx + dy)$$

$$= (\alpha(ax + by) + \beta(cx + dy), \gamma(ax + by) + \delta(cx + dy))$$

$$= (\alpha ax + \alpha by + \beta cx + \beta dy, \gamma ax + \gamma by + \delta cx + \delta dy)$$

$$= ((\alpha a + \beta c)x + (\alpha b + \beta d)y, (\gamma a + \delta c)x + (\gamma b + \delta d)y)$$

Como se cumple la igualdad matricial

$$\begin{pmatrix} \alpha a + \beta c & \alpha b + \beta d \\ \gamma a + \delta c & \gamma b + \delta d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (\alpha a + \beta c)x + (\alpha b + \beta d)y \\ ((\gamma a + \delta c)x + (\gamma b + \delta d)y \end{pmatrix}$$

entonces la matriz asociada a la composición $L_2 \circ L_1$ es

$$C = \begin{pmatrix} \alpha a + \beta c & \alpha b + \beta d \\ \gamma a + \delta c & \gamma b + \delta d \end{pmatrix}$$

Por otro lado, las matrices asociadas a L_1 y L_2 son

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

respectivamente, y un cálculo prueba que el producto

$$BA = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \alpha a + \beta c & \alpha b + \beta d \\ \gamma a + \delta c & \gamma b + d \end{pmatrix} = C$$

lo que nos dice que la composición $L_2 \circ L_1$ tiene asociada al producto de las matrices correspondientes en el mismo orden. $\triangleright$

Generalizamos la anterior discusión en el siguiente resultado.



TEOREMA 5.2 *Para una pareja de transformaciones lineales*

$$L_1 : \mathbb{R}^n \to \mathbb{R}^m, \quad L_2 : \mathbb{R}^m \to \mathbb{R}^k$$

cuyas matrices asociadas son respectivamente

$$A = A_{m\times n} \ (para \quad L_1) \quad y \quad B = B_{k\times m} \ (para \quad L_2)$$

La composición $L_2 \circ L_1$ tiene asociada a la matriz producto

$$C = BA = B_{k\times m}A_{m\times n} = C_{k\times n}$$

EJEMPLO. Sean dadas las transformaciones lineales

$$L_1(x, y) = (2x - y, \ y + x, 4y)$$

$$L_2(x, y, z) = (-x - y - z, \ z + 3x)$$

◁ Las matrices asociadas respectivas son

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 1 \\ 0 & 4 \end{pmatrix} \quad (\text{para } L_1)$$

$$B = \begin{pmatrix} -1 & -1 & -1 \\ 3 & 0 & 1 \end{pmatrix} \quad (\text{para } L_2)$$

y por el teorema 5.3, la matriz de la composición L_2 o L_1 es entonces,

$$C = BA = \begin{pmatrix} -1 & -1 & -1 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 1 & 1 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} -3 & -4 \\ 6 & 1 \end{pmatrix}$$

Por lo tanto, al aplicar convencionalmente en un vector se tiene,

$$(L_2 \circ L_1)(x, y) = \begin{pmatrix} -3 & -4 \\ 6 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -3x - 4y \\ 6x + y \end{pmatrix}$$

lo cual se puede escribir en forma transpuesta como

$$(L_2 \circ L_1)(x, y) = (-3x - 4y, 6x + y) \quad \triangleright$$



5.2 La matriz jacobiana

Comenzamos en esta sección, a discutir la propiedad de diferenciabilidad de una función vectorial de variable vectorial, generalizando lo que se obtuvo para caso de funciones reales de variable vectorial, donde intervenía fuertemente la continuidad de las derivadas parciales que conforman el gradiente. El objeto matemático necesario para generalizar el gradiente viene dado por una matriz cuyas entradas en cada punto son las derivadas parciales de las funciones coordenadas, respecto a todas las variables independientes. Esta matriz lleva el nombre de **Matriz Jacobiana**.

Sea $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ una función tal que en coordenadas del codominio se escriba $f(p) = (f_1(p), \cdots, f_m(p))$, para $p \in \Omega \subset \mathbb{R}^n$.

Cada función coordenada es real de variable vectorial

$$\begin{cases} f_1 : \Omega \to \mathbb{R} \\ f_2 : \Omega \to \mathbb{R} \\ \quad\vdots \\ f_m : \Omega \to \mathbb{R} \end{cases}$$

Si cada función $f_i : \Omega \to \mathbb{R}$ tiene un gradiente en el punto p para toda $1 = 1, 2, \cdots, m$

$$\nabla(f_i)_p = \left(\frac{\partial f_i}{\partial x_1}(p), \quad \frac{\partial f_i}{\partial x_2}(p), \cdots, \frac{\partial f_i}{\partial x_n}(p)\right)$$

entonces podemos formar la siguiente matriz $m \times n$,

$$Df_p = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \frac{\partial f_1}{\partial x_2}(p) \cdots \frac{\partial f_1}{\partial x_n}(p) \\ \frac{\partial f_2}{\partial x_1}(p) & \frac{\partial f_2}{\partial x_2}(p) \cdots \frac{\partial f_2}{\partial x_n}(p) \\ \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \frac{\partial f_m}{\partial x_2}(p) \cdots \frac{\partial f_m}{\partial x_n}(p) \end{pmatrix}_{m\times n}$$

llamada La **matriz Jacobiana** de la función f en el punto p.

Damos ahora algunos ejemplos en dimensiones bajas, esto es, para cuando $n, m \leq 3$.

EJEMPLO. Sea la función $f : \mathbb{R}^2 \to \mathbb{R}^2$ dada por la regla de correspondencia.

$$f(x, y) = (x + y, x^2y)$$

$\triangleleft$ Debido a que las funciones coordenadas son

$$f_1(x, y) = x + y, \quad f_2(x, y) = x^2y$$



La matriz Jacobiana en cualquier $p \in \mathbb{R}^2$ es

$$Df_p = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2xy & x^2 \end{pmatrix}_p$$

De esta manera, en el punto particular del dominio $p = (1, 2)$, se tiene la matriz Jacobiana de 2×2,

$$Df_{(1,2)} = \begin{pmatrix} 1 & 1 \\ 2xy & x^2 \end{pmatrix}_{(1,2)} = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} \quad \triangleright$$

EJEMPLO. Sea la función $f : \mathbb{R}^2 \to \mathbb{R}^3$ dada por la regla,

$$g(x, y) = (xy, \text{sen}x, x^2 + y^2)$$

$\triangleleft$ Ya que las funciones coordenadas son

$$g_1(x, y) = xy, \quad g_2(x, y) = \text{sen}x, \quad g_3(x, y) = x^2 + y^2$$

entonces en cualquier punto p del codominio se tiene una Matriz Jacobiana de 3×2,

$$Dg_p = \begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \\ \frac{\partial g_3}{\partial x} & \frac{\partial g_3}{\partial y} \end{pmatrix}_p = \begin{pmatrix} y & x \\ \cos x & 0 \\ 2x & 2y \end{pmatrix}_p$$

Por ejemplo, si $p = (\pi, \pi/2)$ entonces, al evaluar se obtiene la matriz

$$Dg_{(\pi,\pi/2)} = \begin{pmatrix} y & x \\ \cos x & 0 \\ 2x & 2y \end{pmatrix}_{(\pi,\pi/2)} = \begin{pmatrix} \pi/2 & \pi \\ -1 & 0 \\ 2\pi & \pi \end{pmatrix} \quad \triangleright$$

En este momento damos pauta para iniciar el estudio de la propiedad de diferenciabilidad de una función vectorial de argumento vectorial extendiendo la definición de diferenciabilidad de un campo escalar.

DEFINICIÓN. Sean $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, una función vectorial de argumento vectorial y un punto $p \in \Omega$. Se dice que f es diferenciable en p si existen, una función $g : \mathbb{R}^n \to \mathbb{R}^m$ y una transformación lineal $L_p : \mathbb{R}^n \to \mathbb{R}^m$ tales que

$$\begin{cases} f(p + h) - f(p) = L_p(h) + ||h||g(h) \\ \text{y} \\ \lim_{||h|| \to 0} g(h) = 0 \end{cases}$$

A continuación se propone una condición suficiente y necesaria para la diferenciabilidad de un campo vectorial.

TEOREMA 5.3 (de diferenciabilidad) *Una función $f : \Omega \to \mathbb{R}^m$ es* **diferenciable** *en el punto p, si y sólo sí, La matriz Jacobiana de f en p existe y las entradas*

$$Df_p = \left(\frac{\partial f_i}{\partial x_j}\right)_p, i = 1, 2, \cdots, m \quad j = 1, 2 \cdots, n$$

son continuas,

Esto es, cada función $\frac{\partial f_i}{\partial x_j}$ es continua en p. En este caso, la transformación lineal $L_p : \mathbb{R}^n \to \mathbb{R}^m$ tiene como asociada a La matriz Jacobiana.

DEFINICIÓN. A la transformación lineal

$$L_p : \mathbb{R}^n \to \mathbb{R}^m$$

se le llama la **diferencial** de la función f en el punto p. A su matriz asociada en la base canónica (**Jacobiana**)

$$Df_p = \left(\frac{\partial f_i}{\partial x_j}\right)_p$$

se le llamará la **derivada** de la función f en el punto p.

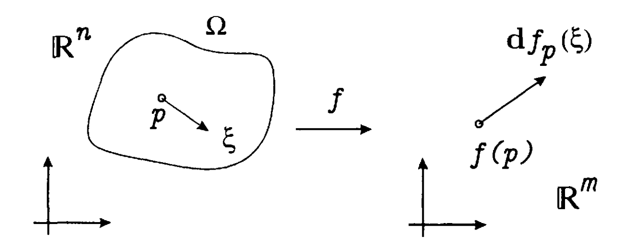


Figura 5.7: Aplicación de un vector ξ bajo la diferencial df_p.

Observemos en realidad que la diferencial L_p

$$L_p : \mathbb{R}^n \to \mathbb{R}^m$$

se aplica en los vectores de posición ξ con extremo en p,

$$\xi \to L_p(\xi)$$

bajo La matriz Jacobiana Df_p.



Entenderemos por $L_p = df_p$ a la transformación lineal asociada a Df_p,

$$df_p : \mathbb{R}^n \to R^m$$

que define una transformación lineal entre los espacios $\mathbb{R}^n$ y $\mathbb{R}^m$. Esto es, si ξ es un vector de $\mathbb{R}^n$ entonces $df_p(\xi)$ es un vector de R^m (véase la figura 5.7).

A una función que es diferenciable en cada punto de su dominio se le llamará **diferenciable**.

EJEMPLO. Sean, la función diferenciable

$$f(x, y) = (x \cos x, \text{ sen } x)$$

el punto $p = \left(\frac{\pi}{2}, \pi\right)$, y el vector $\xi = (-1, 2)$. Calcular $df_p(\xi)$.

$\lhd$ Calculamos La matriz Jacobiana de la función en el punto p,

$$Df_p = \begin{pmatrix} \cos y & -x \text{ sen } y \\ x \cos z & \text{sen } x \end{pmatrix}_p$$

y la evaluamos en el punto $p = \left(\frac{\pi}{2}, \pi\right)$, obteniendo,

$$Df_{(\pi/2,\pi)} = \begin{pmatrix} \cos \pi & -\pi/2 \text{ sen } \pi \\ \pi \cos \pi/2 & \text{sen } \pi/2 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Por lo tanto, para $\xi = (-1, 2)$ se tiene, al aplicar convencionalmente, que

$$df_p(-1, 2) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} = (1, 2) \quad \rhd$$

EJEMPLO. Sean, la función diferenciable,

$$f(x, y) = (\ln xy, x^2 + y^2, \sqrt{xy}),$$

el punto $p = (4, 3)$, y el vector $\xi = (2, 1)$. Calcular $df_p(\xi)$.

$\lhd$ Calculamos La matriz Jacobiana de la función en el punto arbitrario p,

$$Df_p = \begin{pmatrix} \frac{1}{x} & \frac{1}{y} \\ 2x & 2y \\ \frac{1}{2}\sqrt{\frac{y}{x}} & \frac{1}{2}\sqrt{\frac{x}{y}} \end{pmatrix}_p$$



de donde, al evaluar en $p = (4,3)$ se tiene que

$$Df_{(4,3)} = \begin{pmatrix} \frac{1}{4} & \frac{1}{3} \\ 8 & 6 \\ \frac{1}{2}\sqrt{\frac{3}{4}} & \frac{1}{2}\sqrt{\frac{4}{3}} \end{pmatrix}_p = \begin{pmatrix} \frac{1}{4} & \frac{1}{3} \\ 8 & 6 \\ \frac{\sqrt{3}}{4} & \frac{\sqrt{3}}{3} \end{pmatrix}$$

Por lo tanto, al aplicar en el vector $\xi = (2,1)$ de manera convencional, se obtiene,

$$df_{(4,3)}(\xi) = \begin{pmatrix} \frac{1}{4} & \frac{1}{3} \\ 8 & 6 \\ \frac{\sqrt{3}}{4} & \frac{\sqrt{3}}{3} \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{5}{6} \\ 22 \\ \frac{5\sqrt{3}}{6} \end{pmatrix} = \left(\frac{5}{6}, 22, \frac{5\sqrt{3}}{6} \right) \quad \rhd$$

5.3 La regla de la cadena

En el capítulo 4 hemos dado una versión más general del teorema de diferenciabilidad, para la composición de funciones, para cuando eran del tipo $\mathbb{R} \to \mathbb{R}^3 \to \mathbb{R}$. En esta sección damos un resultado más general acerca de la diferenciabilidad de la composición de funciones diferenciables del tipo $\mathbb{R}^n \to \mathbb{R}^m \to \mathbb{R}^k$ (k, n, m enteros arbitrarios), llamada también **Regla de la cadena** (véase la figura 5.8).

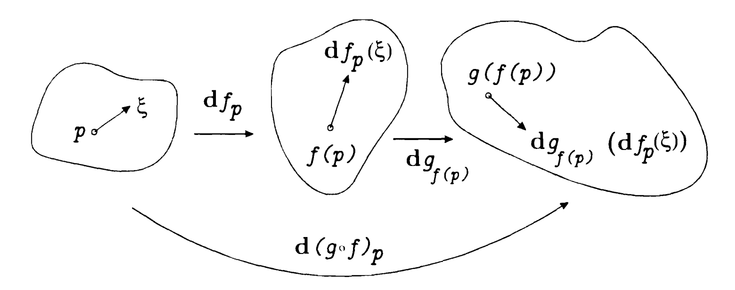


Figura 5.8: Regla de la cadena.

EJEMPLO. Sea la pareja de funciones vectoriales

$$f(x,y) = \left(\frac{y}{1+x^2}, xy + y^2 \right)$$

$$g(u,v,w) = (uv - w,\ u + v + w)$$

$\lhd$ Las matrices Jacobianas respectivas están dadas por,

$$Df_q = \begin{pmatrix} \frac{-2xy}{(1+x^2)^2} & \frac{1}{1+x^2} \\ y & x + 2y \end{pmatrix}_q$$



$$Dg_p = \begin{pmatrix} v & u & -1 \\ 1 & 1 & 1 \end{pmatrix}_p$$

Observamos que sólo tiene sentido la composición dada por

$$f \circ g : \mathbb{R}^3 \to \mathbb{R}^2 \to \mathbb{R}^2$$

la cual ya ha sido calculada y está definida por la regla

$$(f \circ g)(u, v, w) = \left(\frac{u + v + w}{1 + (uv - w)^2}, (u + v + w)(uv + u + v) \right)$$

y cuya matriz Jacobiana se calcula en un punto $p \in \mathbb{R}^3$ por

$$D(f \circ g)_p$$

$$= \begin{pmatrix} \frac{1+(uv-w)(uv+w+2v^2+2vw)}{(1+(uv-w)^2)^2} & \frac{1+(uv-w)(uv+w+2u^2+2uv)}{(1+(uv-w)^2)^2} & \frac{1+(uv-w)(u+uv+2w+2v)}{(1+(uv-w^2)^2} \\ v^2 + 2u + 2v + 2uv + w + w & u^2 + 2uv + 2u + 2v + uw + w & uv + u + v \end{pmatrix}$$

cuyo cálculo resulta ser muy complicado.

Por ejemplo, si se toma el punto $p = (1, 1, 1)$ entonces,

$$D(f \circ g)_{(1,1,1)} = \begin{pmatrix} 1 & 1 & 1 \\ 9 & 9 & 3 \end{pmatrix}$$

Por otro lado, sabiendo que $g(p) = g(1, 1, 1) = (0, 3) = q$ se tendrá

$$Df_q = Df_{(0,3)} = \begin{pmatrix} 0 & 1 \\ 3 & 6 \end{pmatrix}$$

$$Dg_p = Dg_{(1,1,1)} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

Pero al realizar el producto de matrices se tiene que

$$Df_q Dg_p = \begin{pmatrix} 0 & 1 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 9 & 9 & 3 \end{pmatrix} = D(f \circ g)_p$$

lo que nos indica que

$$Df_q Dg_p = D(f \circ g)_p$$



Ya que el punto $q = g(p)$, entonces lo anterior se escribe de manera más explícita y en términos de p como

$$D(f \circ g)_p = Df_{g(p)}\, Dg_p \quad \triangleright$$

La fórmula que se ha obtenido al final del ejemplo será propuesta como **La regla de la cadena**. El enunciado será dado en términos de diferenciales, lo cual no nos causa problema alguno, ya que la prueba involucrará sus matrices Jacobianas asociadas.

TEOREMA 5.4 (Regla de la cadena). *Sean dos funciones diferenciables*

$$f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$$
$$g : U \subset \mathbb{R}^m \to \mathbb{R}^k$$

de tal forma que la imagen de Ω bajo la función f esté contenida en $U \subset \mathbb{R}^m$.

Si $p \in \Omega$ es un punto arbitrario, entonces la función compuesta

$$g \circ f : \Omega \to U \to \mathbb{R}^k$$

es diferenciable en el punto p y la diferencial de tal composición en el punto p cumple la igualdad

$$d(g \circ f)_p = dg_{f(p)} \circ df_p$$

$\triangleleft$ Como las matrices asociadas a las diferenciales son las Jacobianas, será suficiente con demostrar la igualdad

$$D(g \circ f)_p = Dg_{f(p)} Df_p$$

Ya que f es diferenciable en $p \in \Omega$, existe una función $\phi : \mathbb{R}^n \to \mathbb{R}^m$ tal que

$$\begin{cases} f(p+h) - f(p) = Df_p(h) + ||h||\phi(h) \\ \lim_{||h|| \to 0} \phi(h) = 0 \end{cases}$$

Análogamente, como g es diferenciable en el punto $q \in U$, existe una función $\psi : \mathbb{R}^n \to \mathbb{R}^\ell$ tal que

$$\begin{cases} g(p+k) - g(q) = Dg_q(k) + ||k||\varphi(k) \\ \lim_{||h|| \to 0} \varphi(k) = 0 \end{cases}$$

Sea ahora, $q = f(p)$ y considérese el incremento $k = f(p+h) - f(p)$, entonces

$$k = f(p+h) - f(p) = Df_p(h) + ||h||\phi(h)$$

y además

$$g(f(p+h)) = g(f(p)+k) = g(f(p)) + Dg_{f(p)}((Df_p(h) + ||h||\phi(h))$$
$$+||Df_p(h) + ||h||\phi(h)||\psi(Df_p(h) + ||h||\phi(h))$$

Por lo tanto, de la linealidad de $Dg_{f(p)}$ se tiene que

$$g(f(p+h)) - g(f(p)) = Dg_{f(p)}Df_p(h) + \varphi(h)$$

donde

$$\varphi(h) = ||h||Dg_{f(p)}(\phi(h)) + ||Df_p(h) + ||h||\phi(h)||\psi(Df_p(h) + ||h||\phi(h))$$

Como $\lim_{||h||\to 0}\phi(h) = 0$ y $k \to 0 \iff h \to 0$, se tiene que

$$\lim_{||h||\to 0} \varphi(h) = 0$$

sabiendo que $\varphi : \mathbb{R}^n \to \mathbb{R}^\ell$.

De esta manera, se cumple la pareja de igualdades

$$\begin{cases} g(f(p+h)) - g(f(p)) = Dg_{f(p)}Df_p(h) + \varphi(h) \\ \lim_{||h||\to 0}\varphi(h) = 0 \end{cases}$$

o equivalentemente,

$$\begin{cases} (g\circ f)(p+h) - (g\circ f)(p) = Dg_{f(p)}Df_p(h) + \varphi(h) \\ \lim_{||h||\to 0}\varphi(h) = 0 \end{cases}$$

lo que demuestra el teorema. ▷

EJEMPLO. Dada la pareja de funciones,

$$f(x,y,z) = (x^2, y^2 + z^2)$$

$$g(u,v) = (\ln(u+v), e^{uv}, uv)$$

calcular $D(f\circ g)_p$, donde $p = (1,1)$.

◁ Las matrices Jacobianas están dadas por

$$Dg_p = \begin{pmatrix} \frac{1}{u+v} & \frac{1}{u+v} \\ ve^{uv} & ue^{uv} \\ v & u \end{pmatrix}_p$$

$$Df_q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2y & 2z \end{pmatrix}_q$$



Ya que $g(p) = g(1,1) = (\ln 2, e, 1) = q$ se tiene que

$$Dg_{(1,1,1)} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ e & e \\ 1 & 1 \end{pmatrix}$$

$$Df_{(\ln 2,e,1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2e & 2 \end{pmatrix}$$

Por La regla de la cadena, tenemos finalmente,

$$D(f \circ g)_{(1,1)} = Df_{(\ln 2,e,1)} Dg_{(1,1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2e & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ e & e \\ 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1/2 & 1/2 \\ 2e^2 + 2 & 2e^2 + 2 \end{pmatrix} \quad \triangleright$$

EJEMPLO. Sean $\gamma(t) = (x(t), y(t), z(t))$ una curva diferenciable y $w = f(x, y, z)$ una función arbitraria real de variable vectorial.

Si $\gamma(t) \subset \Omega$, siendo Ω el dominio de f en $\mathbb{R}^3$ entonces,

$$w(\gamma(t)) = f(x(t), y(t), z(t))$$

es una función real de variable real diferenciable, $w \circ \gamma : \mathbb{R} \to \mathbb{R}$.

$\triangleleft$ Si $p = \gamma(t)$ es un punto sobre la curva, por La regla de la cadena se tiene

$$\frac{d(w(\gamma(t))}{dt} = D(w \circ \gamma)_t = Dw_{\gamma(t)} D\gamma(t) = Df_{\gamma(t)} D\gamma(t)$$

$$= \left(\frac{\partial f}{\partial x} \ \frac{\partial f}{\partial y} \ \frac{\partial f}{\partial z} \right)_{\gamma(t)} \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \end{pmatrix}_t = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dy} + \frac{\partial f}{\partial z}\frac{dz}{dt}$$

$$= < \nabla f_{\gamma(t)}, \dot{\gamma}(t) >$$

Notamos que esta igualdad es la que se obtiene en el capítulo 4 como un caso especial del teorema de La regla de la cadena establecido en este capítulo. $\triangleright$



5.4 Cambios de coordenadas

En esta sección mostramos uno de los resultados más importantes del análisis matemático moderno, el teorema de la **Función Inversa**. Aunque se omite su demostración, se ilustra su alcance con ejemplos y se dan referencias para los lectores interesados en la parte teórica.

EJEMPLO. Consideremos el siguiente sistema lineal de ecuaciones de 3×3,

$$\begin{cases} x + 2y + 3z = 1 \\ 2x + 5y + 3z = 2 \\ x \quad\; + 8z = 3 \end{cases}$$

$\triangleleft$ Si se consiera la matriz del sistema,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix}$$

se tiene que $detA = -1$, de donde A es invertible con inversa,

$$A^{-1} = \begin{pmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix}$$

que puede ser calculada mediante El método de Gauss-Jordan.

Ya que el sistema dado se puede escribir matricialmente

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \Longleftrightarrow A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

entonces el vector columna con las incógnitas puede despejarse mediante la multiplicación de A^{-1}, esto es,

$$A^{-1}A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \Longleftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

lo que nos indica que

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -40 & -16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ -6 \\ -2 \end{pmatrix}$$

y, por lo tanto, que el punto con coordenadas $(x, y, z) = (3, -6 - 2)$ es la solución a tal sistema.

Veamos el problema inicial desde otro ángulo.

Si consideremos a $L_A : \mathbb{R}^2 \to \mathbb{R}^3$ la transformación lineal asociada a la matriz A, entonces el sistema de ecuaciones inicial puede ser visto mediante la ecuación funcional (con notación convencional),

$$L_A(x, y, z) = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = (1, 2, 3)$$

es decir, el problema consiste en encontrar un punto (x, y, z) del dominio de L_A tal que bajo la aplicación de la transformación tenga imagen en el punto dado $(1, 2, 3)$,

$$L_A(x, y, z) = (1, 2, 3)$$

Por el proceso mostrado, se tiene que si $detA \neq 0$, y $L_{A^{-1}}$ es la transformación lineal asociada a la matriz A^{-1}, entonces el punto buscado es

$$(x, y, z) = L_{A^{-1}}(1, 2, 3) = (3, -3, -2).$$

Si se analiza un sistema de ecuaciones más general,

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = u \\ a_{21}x + a_{22}y + a_{23}z = v \\ a_{31}x + a_{32}y + a_{33}z = w \end{cases}$$

con la matriz asociada al sistema dada por

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

entonces, una condición para que el sistema tenga solución para las incógnitas (x, y, z) en términos de los números (u, v, w) es que $detA \neq 0$. Si éste es el caso y se define por

$$A^{-1} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

entonces la solución al sistema de ecuaciones viene dado por la ecuación matricial

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$



En términos funcionales, si $L_A : \mathbb{R}^3 \to \mathbb{R}^3$ es la transformación lineal asociada a la matriz A y $L_{A^{-1}} : \mathbb{R}^3 \to \mathbb{R}^3$ es la transformación lineal asociada a la matriz A^{-1}, el sistema de ecuaciones en las incógnitas (x, y, z) y su solución en términos de (u, v, w) vienen dados por la equivalencia funcional

$$L_A(x, y, z) = (u, v, w) \Longleftrightarrow (x, y, z) = L_{A^{-1}}(u, v, w)$$

Esta equivalencia señala explícitamente las propiedades de invertibilidad de cada transformación lineal mencionada.

De hecho, de la igualdad $AA^{-1} = A^{-1}A = I$, donde I es la matriz identidad se tiene que

$$L_A \circ L_{A^{-1}} = L_{AA^{-1}} = L_I = I$$

$$L_{A^{-1}} \circ L_A = L_{A^{-1}A} = L_I = I$$

lo que nos dice que L_A y $L_{A^{-1}}$ son mutuamente inversas y que

$$L_{A^{-1}} = (L_A)^{-1} \quad \triangleright$$

Este problema de invertibilidad se puede formular en términos más generales, cuando se tiene una función vectorial de variable vectorial.

Como se entiende del ejemplo anterior, las funciones inversas sirven como operaciones de cancelamiento o despeje (véase Capítulo I de **Reyes**, 1996).

DEFINICIÓN. Sea $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ una función vectorial de variable vectorial. La función $g : f(\Omega) \subset \mathbb{R}^n \to \Omega$ se dice ser una **función inversa** de la función si se cumple que

$$f \circ g : f(\Omega) \to \Omega \to f(\Omega), \quad \text{es tal que} \quad f \circ g = 1$$

$$g \circ f : \Omega \to f(\Omega) \to \Omega, \quad \text{es tal que} \quad g \circ f = 1$$

siendo I la función identidad (véase la figura 5.9).

EJEMPLO. Sea la función $f : \mathbb{R}^2 \to \mathbb{R}^2$ dada por la regla de correspondencia

$$f(x, y) = (e^x \cos y,\ e^x \operatorname{sen} y)$$

$\triangleleft$ Si se escriben las coordenadas de la imagen,

$$\begin{cases} u = e^x \cos y \\ v = e^x \operatorname{sen} y \end{cases}$$



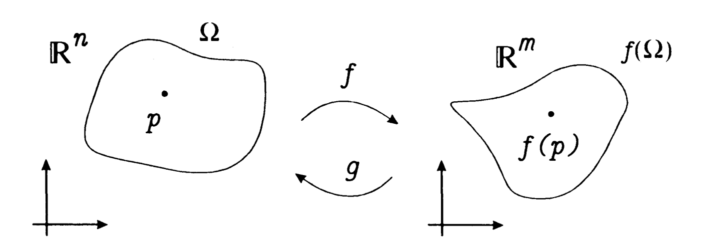


Figura 5.9: Invertibilidad de una función.

entonces, al dividir la segunda ecuación por la primera se obtiene

$$\frac{v}{u} = \tan y \Rightarrow y = \arctan\left(\frac{u}{v}\right)$$

Si se elevan al cuadrado las igualdades en ambos miembros y se suman los lados respectivos se tendría que

$$u^2 + v^2 = e^{2x}\cos^2 y + e^{2x}\operatorname{sen}^2 y = e^{2x}$$

lo que implica que

$$x = \ln\sqrt{u^2 + v^2}$$

De esta manera, con base en un cálculo realizado mediante despejes, se obtiene una función

$$(x, y) = g(u, v) = \left(\ln\sqrt{u^2 + v^2}, \arctan\left(\frac{v}{u}\right)\right)$$

la cual tiene un dominio de definición

$$D_1 = \mathbb{R}^2 - \{(u, v) | \, u \neq 0\} = \mathbb{R}^2 - \{\text{eje} \quad v\}$$

En otras palabras, bajo esta regla g obtenida, no todo punto $(u, v) \in \mathbb{R}^2$ acepta solución para la equivalencia

$$f(x, y) = (u, v) \iff (x, y) = g(u, v).$$

Por eso es que llamamos a g una función inversa, porque está definida no en toda la imagen de f.

Otro cálculo simple nos muestra que la función.

$$h(u, v) = \left(\ln\sqrt{u^2 + v^2}, \operatorname{arccot}\left(\frac{u}{v}\right)\right)$$

es una función obtenida mediante el método similar, y que es inversa para f definida en el conjunto

$$D_2 = \mathbb{R}^2 - \{\text{eje}\, u\}$$

Lo que es claro, es que un cálculo directo nos mostraría que en el conjunto

$$D_1 \cap D_2 = D = \mathbb{R}^2 - \{\text{ejes}\, u, v\}$$

las funciones h y g coinciden en sus valores. $\triangleright$

EJEMPLO. Sea $f : D \subset \mathbb{R} \to \mathbb{R}$ una función diferenciable real de variable real y sea $p \in D$ un **punto regular** de f, es decir la derivada $f'(p)$ no se anula.

$\triangleleft$ Por el teorema de la función inversa (Capítulo 3 Reyes, 1996) se tiene que existen intervalos J_p alrededor de p y J_q alrededor de $q = f(p)$ de tal forma que la restricción de la función f a esos intervalos

$$f : J_p \to J_q$$

es invertible y su inversa es diferenciable. Más aún, si $g : J_q \to J_p$ es la función inversa local de f alrededor de $q = f(p)$, entonces la ecuación

$$(g \circ f)(x) = x$$

implica, por La regla de la cadena

$$1 = (g \circ f)'(x) = g'(f(x))f'(x)$$

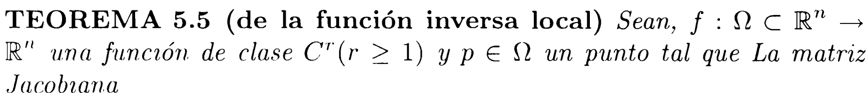
Al evaluar en $x = p$ se tiene que la derivada de la función inversa g en el punto $q = f(p)$ se calcula,

$$g'(q) = (f'(p))^{-1} = \frac{1}{f'(p)}$$

Decimos en este caso que la función f es una **invertible local** pues tiene una inversa local g que también es diferenciable. $\triangleright$

El siguiente resultado generaliza el ejemplo anterior y nos provee de una gran cantidad de funciones localmente invertibles de clase C^r. La prueba se omite, pero el lector interesado puede referirse al trabajo de **Lima** (1992).

Para una función de clase C^r, $f : \mathbb{R}^n \to \mathbb{R}^m$ entenderemos a una función cuyas funciones coordenadas $f_i : \mathbb{R}^n \to \mathbb{R}`m$ son de clase C^r.



TEOREMA 5.5 (de la función inversa local) *Sean, $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ una función de clase $C^r (r \geq 1)$ y $p \in \Omega$ un punto tal que La matriz Jacobiana*

$$A = Df_p = \left(\frac{\partial f_i}{\partial x_j}\right)_p$$

tenga un determinante no nulo: $detA \neq 0$. Entonces f es una función localmente invertible en p. Esto es, en una vecindad de p y en otra vecindad de $q = f(p)$. la función tiene una inversa local g que es de clase C^r.

Más aún, si $q = f(p)$ es la imagen se tiene que La matriz Jacobiana de g en q se calcula invirtiendo la matriz Df_p, es decir,

$$Dg_q = (Df_p)^{-1}$$

EJEMPLO. Considerar nuevamente la función $f : \mathbb{R}^2 \to \mathbb{R}^2$

$$f(x, y) = (e^x \cos y,\ e^x \operatorname{sen} y)$$

de la cual no se puede obtener una función inversa global explícita.

$\lhd$ Sea $p = (x, y) \in \mathbb{R}^2$ un punto arbitrario del dominio. Su matriz Jacobiana viene dada por

$$Df_p = \begin{pmatrix} e^x \cos y & -e^x \operatorname{sen} y \\ e^x \operatorname{sen} y & e^x \cos y \end{pmatrix}$$

lo que nos dice que

$$det(Df_p) = \begin{vmatrix} e^x \cos y & -e^x \operatorname{sen} y \\ e^x \operatorname{sen} y & e^x \cos y \end{vmatrix} = e^x \neq 0$$

Por lo tanto, la función f es invertible localmente en cada punto y dependiendo de éste, tendrá una función inversa local como alguna de las dos diferentes halladas anteriormente, según la forma de despejar las variables. $\rhd$

DEFINICIÓN. Dada La matriz Jacobiana en un punto p de una función $f : \mathbb{R}^n \to \mathbb{R}^n$ de clase C^r, definida por Df_p, a su determinante $det(Df_p)$ se le llama el **Jacobiano** de la función f en el punto p y lo calculamos por Jf_p, es decir

$$Jf_p = det(Df_p)$$

Así, repasamos el teorema de la función inversa diciendo que una función $\mathbb{R}^n \to \mathbb{R}^n$ de clase C^r cuyos Jacobianos no se anulan en todo el dominio es localmente invertible de clase C^r.

Para optimizar cálculos utilizaremos la notación clásica que utiliza subíndices variables para definir la derivación parcial. Por ejemplo,

$$f_u = \frac{\partial f}{\partial u}, \quad f_x = \frac{\partial f}{\partial x}, \quad u_z = \frac{\partial u}{\partial z}, \cdots$$

definirá la derivación parcial respecto a la variable indicada.

Por otro lado, si una función $\mathbb{R}^n \to \mathbb{R}^n$ es definida por

$$(y_1 \cdots, y_n) = f(x_1 \cdots, x_n)$$

el jacobiano de f en el punto p se definirá por

$$Jf_p = \det\left(\frac{\partial y_i}{\partial x_j}\right)_p = \frac{D(y_1, \cdots, y_n)}{D(x_1 \cdots, x_n)}(p)$$

EJEMPLO. En la función suave $\mathbb{R}^2 \to \mathbb{R}^2$ dada por

$$(u, v) = f(x, y) = (e^x \cos y, e^x \operatorname{sen} y)$$

se tiene

$$\lhd \quad u_x = \frac{\partial u}{\partial x} = e^x \cos y, \quad u_y = \frac{\partial u}{\partial y} = -e^x \operatorname{sen} y$$

$$v_x = \frac{\partial v}{\partial x} = e^x \operatorname{sen} y, \quad v_y = \frac{\partial v}{\partial y} = e^x \cos y$$

$$Jf_p = \frac{D(u,v)}{D(x,y)} = \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} = \begin{vmatrix} e^x \cos y & -e^x \operatorname{sen} y \\ e^x \operatorname{sen} y & e^x \cos y \end{vmatrix} = e^x \neq 0$$

Particularmente, en el punto $p = (1, 1)$ se tiene que

$$Jf_{(1,1)} = e^1 = e$$

de donde f es invertible alrededor del punto $(1, 1)$ con una inversa local

$$g(u, v) = \left(\ln \sqrt{u^2 + v^2}, \arctan\left(\frac{u}{v}\right)\right)$$

definida alrededor del punto $f(1, 1) = (e \cos 1,\ e \operatorname{sen} 1)$. $\rhd$

EJEMPLO. Dada la función

$$f(x, y, z) = (xz, xy, yz)$$

verificar su invertibilidad alrededor del punto $p = (-1, 1, -1)$.



$\lhd$ Si calculamos por $(u, v, w) = f(x, y, z) = (xz, xy, yz)$, es decir,

$$\begin{cases} u = xz \\ v = xy \\ w = yz \end{cases}$$

entonces

$$Jf_p = \frac{D(u, v, w)}{D(x, y, z)} = \begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} = \begin{vmatrix} z & 0 & x \\ y & x & 0 \\ 0 & z & y \end{vmatrix} = 2xyz$$

De esta manera, el Jacobiano se anula sólo en los planos coordenados $x = y = z = 0$.

Por lo tanto, $Jf_{(1,1,1)} = 2 \neq 0$, lo que nos dice que f es invertible en una localidad del punto $(1, 1, 1)$, y procedemos a calcular en este caso una función inversa local de f.

De las ecuaciones $u = xz$, $v = xy$ se obtiene que $uv = x^2zy$. Al utilizar ahora $w = yz$ en $uv = x^2zy$ se tiene que $uv = x^2w$ lo que nos dice que

$$x^2 = \frac{uv}{w} \iff x = \pm\sqrt{\frac{uv}{w}}$$

Análogamente, se obtienen,

$$y = \pm\sqrt{\frac{vw}{u}}, \; z = \pm\sqrt{\frac{uw}{v}}.$$

Ya que $q = (1, -1, -1) = f(-1, 1, -1)$ está en el dominio para cualquier elección de signos, se tiene, al elegir los apropiados, que la inversa local g tiene la forma,

$$g(u, v, w) = \left(-\sqrt{\frac{vv}{w}}, \sqrt{\frac{vw}{u}}, -\sqrt{\frac{uw}{v}}\right)$$

en virtud de los signos del punto $p = (-1, 1, -1)$. $\rhd$

En muchas ocasiones, el sistema cartesiano de coordenadas no es bueno para describir objetos geométricos o mecánicos que están incluidos en el espacio vectorial $\mathbb{R}^3$. Por eso, es preciso utilizar otro tipo de coordenadas que describen con fórmulas más simples tales objetos. De esta forma, si se conocen las coordenadas que se van a utilizar en lugar de las cartesianas,

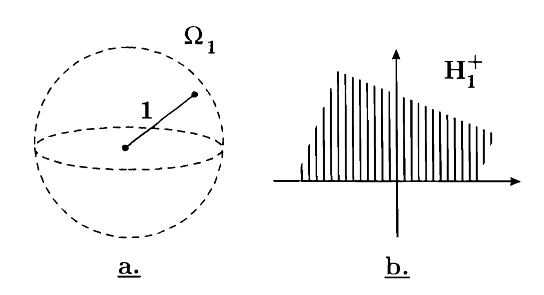


Figura 5.10: Regiones en $\mathbb{R}^3$ y $\mathbb{R}^2$.

justo es que se establezca una relación biunívoca entre las colecciones correspondientes a tales sistemas de coordenadas.

Cabe mencionar que en muchas ocasiones, tal relación biunívoca entre tales sistemas puede establecerse entre **regiones.**

EJEMPLO. Sea el campo escalar

$$F(x, y, z) = x^2 + y^2 + z^2$$

Ya que F es una función suave (por lo tanto, continua) se tiene que

$$\Omega_1 = F^{-1}(-\infty, 1) = \{(x, y, z)| \quad x^2 + y^2 + z^2 < 1\}$$

es una región abierta de $\mathbb{R}^3$ Esto, es la bola de radio 1 con centro en $(0, 0, 0)$ de $\mathbb{R}^3$ es una región abierta (véase figura 5.10 a).

EJEMPLO. Considere el semiplano superior

$$H^+ = \{(x, y) \in \mathbb{R}^2| \quad y > 0\}$$

Este conjunto es una región abierta en $\mathbb{R}^2$, pues la función $F(x, y) = y$ es continua y $H^+ = F^{-1}(0, \infty)$ (véase figura 5.10 b). Claramente la cerradura del semiplano superior

$$\bar{H}^+ = \{(x, y) \in \mathbb{R}^2| \quad y \geq 0\}$$

es una región con frontera en el plano $\mathbb{R}^2$.

En algunas ocasiones, una región está definida por varias desigualdades. Esto es, cada punto de la región satisface un número finito de desigualdades del tipo

$$\begin{cases} F_1(x, y, z) \leq a_1 \\ F_2(x, y, z) \leq a_2 \\ \quad\vdots \\ F_n(x, y, z) \leq a_n \end{cases}$$



donde las funciones $F_1, \cdots, F_n$ son continuas en el espacio $\mathbb{R}^3$, y $a_1, \cdots, a_n$ son números reales.

EJEMPLO. Consideremos la región en el plano $\mathbb{R}^2$ acotada por la elipse

$$\frac{x^2}{4} + \frac{y^2}{2} = 1$$

$\lhd$ De la figura 5.11 se puede observar que la variable x está necesariamente en el intervalo $[-2, 2]$ y que la variable y está encerrada desde la curva $y = \frac{-1}{\sqrt{2}}\sqrt{4-x^2}$ hasta la curva $y = \frac{1}{\sqrt{2}}\sqrt{4-x^2}$.

La definición de y se obtiene al despejar a y de la ecuación inicial,

$$y = \pm\frac{1}{\sqrt{2}}\sqrt{4-x^2}$$

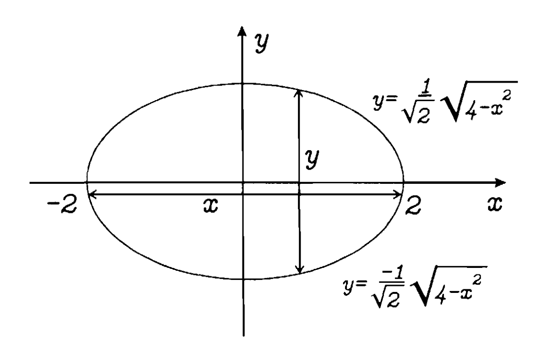


Figura 5.11: Región elíptica en $\mathbb{R}^2$.

De esta forma, la región Ω se puede describir totalmente por las desigualdades

$$-2 \leq x \leq 2 \quad \text{y} \quad -\frac{1}{\sqrt{2}}\sqrt{4-x^2} \leq y \leq \frac{1}{\sqrt{2}}\sqrt{4-x^2}.$$

Estas desigualdades son equivalentes al sistema de desigualdades

$$\left\{ \begin{array}{c} F_1(x,y) = -x \leq 2 \\ F_2(x,y) = x \leq 2 \\ F_3(x,y) = -\left(y + \frac{\sqrt{4-x^2}}{\sqrt{2}}\right) \leq 0 \\ F_4(x,y) = y - \frac{\sqrt{4-x^2}}{\sqrt{2}} \leq 0 \end{array} \right.$$

que provienen del par de desigualdades obtenido. $\rhd$

En general, entender un sistema de desigualdades como el que se obtuvo en la parte final del ejemplo anterior es díficil, y es más cómodo describir a una región por desigualdades como las que se obtuvieron directamente en el ejemplo anterior, las cuales señalan los conjuntos donde era permisible tomar cada una de las variables.

EJEMPLO. Considérese la región Ω en el plano definida por desigualdades

$$y \geq x^2, \quad y \leq 4 - x^2$$

$\triangleleft$ La figura 5.12 nos muestra que la primer desigualdad se refiere a la región por encima de la parábola $y = x^2$, mientras que la segunda desigualdad se refiere a la región por debajo de la parábola $y = 4 - x^2$. Por lo tanto, la región Ω es aquella acotada inferiormente por $y = x^2$ y acotada superiormente por $y = 4 - x^2$.

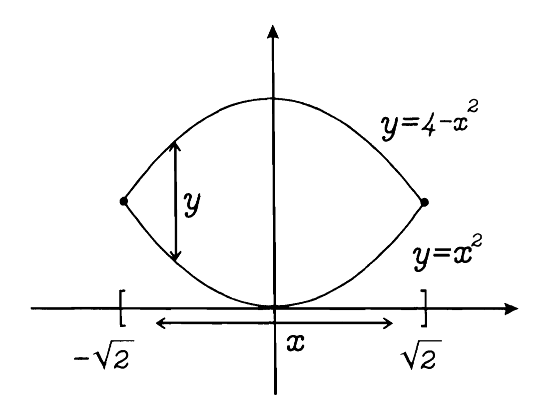


Figura 5.12: Región Ω del plano $\mathbb{R}^2$.

Los puntos de intersección entre las parábolas satisfacen las ecuaciones

$$\begin{cases} y = x^2 \\ y = 4 - x^2 \end{cases}$$

que al igualar: $x^2 = 4 - x^2$, nos da los puntos $x = \pm\sqrt{2}$.

Esto hace que x esté en el intervalo $[-\sqrt{2}, \sqrt{2}]$, mientras que la variable y es permisible entre las curvas $y = x^2$ hasta $y = 4 - x^2$.

Por lo tanto, la región Ω está completamente descrita por las desigualdades

$$-\sqrt{2} \leq x \leq \sqrt{2}, \quad x^2 \leq y \leq 4 - x^2. \quad \triangleright$$

EJEMPLO. Sea $V \subset \mathbb{R}^3$ la región espacial acotada por los planos

$$x = 0, \quad y = 0, \quad z = 0, \quad x + y + z = 1$$



◁ Una inspección a la ecuación del plano $x+y+z=1$, nos muestra que los puntos de intersección con los ejes coordenados son $(1,0,0)$, $(0,1,0)$ y $(0,0,1)$, como se muestra en la figura 5.13, quedando tal porción del plano en el primer octante: $x \geq 0$, $y \geq 0$, $z \geq 0$

Si en la ecuación $x+y+z=1$, ponemos $y=0$ y $z=0$, entonces $x=1$, lo que determina para x el intervalo $[0,1]$.

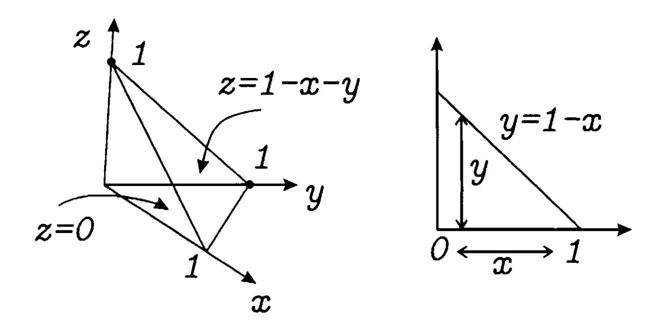


Figura 5.13: Región espacial acotada por planos.

Para determinar el dominio de y, hacemos $z=0$ en la ecuación $x+y+z=1$, obteniendo $x+y=1$ o $y=1-x$, lo que según se muestra en la misma figura, e indica que

$$0 \leq y \leq 1-x$$

Finalmente, al despejar z de la ecuación $x+y+z=1$, se obtiene $z=1-x-y$, lo que dice que la variable z debe satisfacer la desigualdad

$$0 \leq z \leq 1-x-y$$

Por lo tanto, la región espacial V se describe por el conjunto

$$V = \{(x,y,z) | 0 \leq x \leq 1, 0 \leq y \leq 1-x, 0 \leq z \leq 1-x-y\}. \quad \triangleright$$

EJEMPLO. Consideremos la región espacial acotada por la parte superior de la esfera

$$x^2+y^2+z^2=4.$$

◁ Al proceder por analogía al ejemplo anterior haciendo $y=0$, $z=0$ en la ecuación dada, se obtiene que $x^2=4$ o $x=\pm 2$. Al observar la figura 5.14, tenemos que el intervalo numérico para x es $[-2,2]$.



Por otra parte, si se anula z en la ecuación, se tiene que $x^2 + y^2 = 4$ lo que al despejar a y en términos de x nos da $y = \pm\sqrt{4 - x^2}$, que según la figura 5.17, nos indica el dominio para la variable y,

$$-\sqrt{4 - x^2} \leq y \leq \sqrt{4 - x^2}$$

Finalmente, si en la ecuación inicial se despeja a z, se obtiene que $z = \pm\sqrt{4 - x^2 - y^2}$, lo que según la misma figura 5.17, la variable z deberá satisfacer la desigualdad

$$0 \leq z \leq \sqrt{4 - x^2 - y^2}$$

pues se considera la parte superior de la esfera.

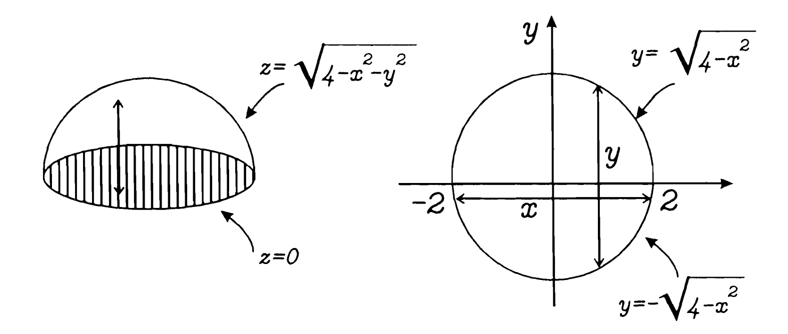


Figura 5.14: Semiesfera superior $\mathbb{R}^3$.

De esta manera, la región $V \subset \mathbb{R}^3$ acotada de esta forma, está totalmente descrita por las desigualdades

$$-2 \leq x \leq 2, \quad -\sqrt{4 - x^2} \leq y \leq \sqrt{4 - x^2}, \quad 0 \leq z \leq \sqrt{4 - x^2 - y^2}\} \quad \triangleright$$

Como se mencionó antes, situaremos un cambio de sistemas coordenados en regiones contenidas en $\mathbb{R}^3$ (ó en $\mathbb{R}^2$).

DEFINICIÓN. Sean x, y, z y u, v, w dos sistemas de coordenadas en $\mathbb{R}^3$. Por un **cambio de sistemas de coordenadas** en una región Ω se entiende una función diferenciable e invertible $\Omega \to \Omega$, que en coordenadas se escribe

$$\begin{cases} x = x(u, v, w) \\ y = y(u, v, w) \\ z = z(u, v, w) \end{cases}$$



con una función inversa diferenciable definida en Ω por

$$\begin{cases} u = u(x, y, z) \\ v = v(x, y, z) \\ w = w(x, y, z) \end{cases}$$

Para $q_0 = (u_0, v_0, w_0) \subset \Omega$ un punto definido en las coordenadas, u, v, w, considérese el correspondiente en Ω en las coordenadas x, y, z, definido por $p_0 = (x_0, y_0, z_0)$.

DEFINICIÓN. El punto $p_0 \in \Omega$ se llama un **valor regular** para el cambio de coordenadas, sí la matriz

$$A = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix}_{q_0}$$

es no singular, es decir, det $A \neq 0$. El punto q_0 se llamará **punto regular**.

En otras palabras, p_0 es un valor regular para el sistema de coordenadas (u, v, w) si la parte lineal A es una matriz invertible.

El siguiente resultado nos habla de la invertibilidad local alrededor de un punto regular de un cambio de coordenadas. Éste se sigue del Teorema de la función inversa.

TEOREMA 5.6 *Si p_0 es un punto regular para el cambio de coordenadas, entonces en una vecindad de p_0 en Ω se puede definir una transformación inversa.*

En otras palabras, la parte lineal del cambio de coordenadas en q_0 define su invertibilidad localmente. Esto es, define localmente un cambio de coordenadas.

A continuación damos ejemplos de cambios de coordenadas en el plano $\mathbb{R}^2$ y en el espacio $\mathbb{R}^3$. Se hará uso del Teorema 5.6 en cada uno de ellos.

Coordenadas polares

Considérese en el plano $\mathbb{R}^2$ el sistema de coordenadas cartesianas (x, y). Introducimos ahora en el mismo plano otro sistema de coordenadas (r, φ), donde $r \geq 0$ y $\varphi \in (0, 2\pi)$. La relación entre los sistemas de coordenadas es

$$\begin{cases} x = r\cos\varphi \\ y = r\,\mathrm{sen}\varphi \end{cases}$$

◁ Dado un punto $(r_0, \varphi_0) = q_0$, procedemos a verificar la condición de invertibilidad local, y para ello necesitamos calcular la matriz Jacobiana en q_0

$$A = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix}_{q_0} = \begin{pmatrix} \cos\varphi_0 & -r_0 \operatorname{sen}\varphi_0 \\ \operatorname{sen}\varphi_0 & r_0 \cos\varphi_0 \end{pmatrix}.$$

De esta forma, el determinante de la parte lineal A se calcula

$$\frac{D(x,y)}{D(r,\varphi)} = detA = \begin{vmatrix} \cos\varphi_0 & -r_0 \operatorname{sen}\varphi_0 \\ \operatorname{sen}\varphi_0 & r_0 \cos\varphi_0 \end{vmatrix} = r_0(\cos^2\varphi_0 + \operatorname{sen}^2\varphi_0) = r_0,$$

con lo que, $detA \neq 0$ en $\mathbb{R}^+ \times (0, 2\pi)$.

Una función que da un cambio inverso de coordenadas está dado por

$$\begin{cases} r = \sqrt{x^2 + y^2} \\ \varphi = \arctan(y/x) \end{cases}$$

que no está definido cuando $x = 0$. De esta manera, ya que la imagen corresponde al eje y (donde $\varphi = \frac{\pi}{2}, \frac{3}{2}\pi$), se tiene una correspondencia de coordenadas (x, y) y (r, φ) en apenas los cuadrantes (sin los ejes) del plano. La figura 5.15 ilustra esto. Al sistema de coordenadas (r, φ) se le llama de **coordenadas polares.** ▷

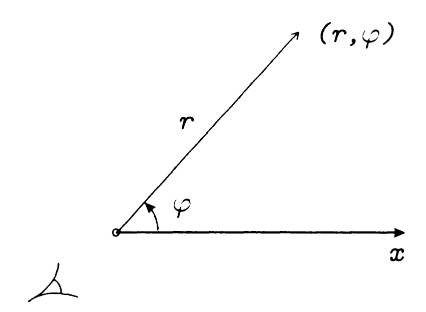


Figura 5.15: Coordenadas Polares en $\mathbb{R}^2$.

EJEMPLO. Consideremos la curva en el plano definida por la ecuación en coordenadas cartesianas

$$25x^2 + 10xy + y^2 - 1 = 0$$

◁ Al realizar el cambio de coordenadas cartesianas por polares, tenemos que la ecuación anterior se transforma en

$$0 = 25x^2 + 10xy + y^2 - 1 = 25(r\cos\varphi)^2 + 10r\cos\varphi\, r\operatorname{sen}\varphi$$



$$+(r\,\mathrm{sen}\,\varphi)^2 - 1 = 25r^2\cos^2\varphi + 10r^2\cos\varphi\,\mathrm{sen}\,\varphi + r^2\mathrm{sen}^2\varphi - 1$$
$$= r^2[25\cos^2\varphi + 10\cos\varphi + \mathrm{sen}^2\varphi] - 1$$
$$= r^2[5\cos\varphi + \mathrm{sen}\,\varphi]^2 - 1$$

o bien, en la relación bivalente

$$r = \frac{\pm 1}{5\cos\varphi + \mathrm{sen}\varphi} \quad \triangleright$$

EJEMPLO. Consideremos la relación en coordenadas polares dada por

$$r = 1 - \cos\varphi$$

$\triangleleft$ Para pasar tal relación a coordenadas cartesianas, procedemos como sigue, $r = 1 - \cos\varphi \iff 1 - r = \cos\varphi$

Al multiplicar por r de cada miembro, se tiene

$$r - r^2 = r\cos\varphi = x$$

o bien,

$$\sqrt{x^2 + y^2} = x + x^2 + y^2$$

Consecuentemente (al elevar al cuadrado a cada lado), obtenemos la relación en coordenadas cartesianas

$$x^4 + y^4 + 2x^3 + 2xy^2 + 2x^2y^2 - y^2 = 0 \quad \triangleright$$

EJEMPLO. Consideremos la función real definida por

$$F(x, y) = \sqrt{x^2 + y^2 - 1}$$

en la región $\Omega = \{(x, y) |\, x^2 + y^2 \geq 1\}$.

$\triangleleft$ Al realizar la transformación de coordenadas cartesianas a polares (r, φ) se tiene una nueva función

$$\widetilde{F}(r, \varphi) = F(r\cos\varphi, r\,\mathrm{sen}\,\varphi) = \sqrt{(r\cos\varphi)^2 + (r\,\mathrm{sen}\,\varphi)^2 - 1} = \sqrt{r^2 - 1}$$

definida en la misma región, en coordenadas polares,

$$\Omega = \{(r, \varphi) |\, r^2 \geq 1,\ \varphi \in [0, 2\pi]\} \quad \triangleright$$

EJEMPLO. Sea Ω la región con frontera en el plano, definida por las coordenadas cartesianas

$$\Omega = \{(x, y) |\, 1 \leq x^2 + y^2 \leq 4\}$$



Este conjunto corresponde a una región anular con centro en $(0,0)$ de radio interior 1 y radio exterior 2.

$\lhd$ Al describir a Ω con coordenadas polares, tenemos

$$\Omega = \{(r,\varphi) \mid 1 \leq r \leq 2,\ \varphi \in [0,2\pi]\}$$

pues la relación dada por

$$1 \leq x^2 + y^2 \leq 4$$

es equivalente en polares a $1 \leq r^2 \leq 4$, ó bien a $1 \leq r \leq 2$. $\rhd$

Coordenadas cilíndricas

Considérese en $\mathbb{R}^3$ el sistema de coordenadas cartesianas (x,y,z) e introdúzcase el sistema nuevo dado por la tripleta (r,φ,z) donde r,φ son las coordenadas del ejemplo anterior, y z es la tercera coordenada cartesiana, esto es,

$$\begin{cases} x = r\cos\varphi, \\ y = r\operatorname{sen}\varphi, \\ z = z \end{cases}$$

Al buscar una condición de invertibilidad para este cambio, se tiene que para cualquier $(r,\varphi,z) \in \mathbb{R}^3$ la matriz Jacobiana del cambio de coordenadas es

$$A = \begin{pmatrix} \cos\varphi & -r\operatorname{sen}\varphi & 0 \\ \operatorname{sen}\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

De igual forma que en el ejemplo anterior, el determinante de A se calcula

$$\frac{D(x,y,z)}{D(r,\varphi,z)} = detA = \begin{vmatrix} \cos\varphi & -r\operatorname{sen}\varphi & 0 \\ \operatorname{sen}\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} \cos\varphi & -r\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & r\cos\varphi \end{vmatrix} = r$$

y este es diferente de cero, si $r > 0$. Este conjunto corresponde a todo el espacio con excepción del eje z. Una función inversa es calculada por

$$\begin{cases} r = \sqrt{x^2+y^2} \\ \varphi = \arctan(y/x) \\ z = z \end{cases}$$

definida cuando $x \neq 0$. Como $\varphi \in (0,2\pi)$, entonces se tiene un cambio de coordenadas (x,y,z) por (r,φ,z), y viceversa, sólo cuando el punto es tomado fuera de los planos $x=0$, $y=0$ en $\mathbb{R}^3$.



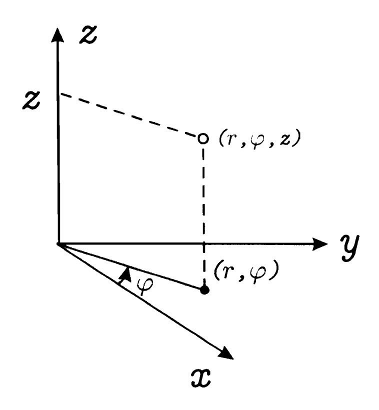


Figura 5.16: Coordenadas cilíndricas.

Este sistema se conoce como **coordenadas cilíndricas.** La figura 5.16 ilustra este

EJEMPLO. Considérese la relación dada en coordenadas cartesianas por

$$x^2 + y^2 - 2z^2 = 0$$

◁ En coordenadas cilíndricas (r, φ, z) esta relación se escribe

$$0 = (x^2 + y^2) - 2z^2 = r^2 - 2z^2,$$

es decir, $r^2 = 2z^2$ o bien, $r = \pm\sqrt{2}z$ ▷

EJEMPLO. Consideremos la función definida por

$$F(x, y, z) = \sqrt{z - x^2 - y^2}$$

en la región $\Omega = \{(x, y, z) |\, z \geq x^2 + y^2\} \subset \mathbb{R}^3$.

◁ En coordenadas cilíndricas la función se escribe de forma más simple

$$\widetilde{F}(r, \varphi, z) = F(r \cos\varphi, r \operatorname{sen}\varphi, z) = \sqrt{z - (r\cos\varphi)^2 - (r\operatorname{sen}\varphi)^2} = \sqrt{z - r^2}$$

definida en la región Ω que en coordenadas cilíndricas es escrita como

$$\Omega = \{(r, \varphi, z) \mid z \geq r^2,\ \varphi \in [0, 2\pi]\}. \quad ▷$$

EJEMPLO. La región espacial anular-cilíndrica dada por

$$\Omega = \{(x, y, z) |\, 1 \leq x^2 + y^2 \leq 4,\ 0 \leq z \leq 1\}$$

en coordenadas cilíndricas se escribe

$$\Omega = \{(r, \varphi, z) |\, 1 \leq r \leq 2,\ 0 \leq z \leq 1,\ \varphi \in [0, 2\pi]\}.$$



Coordenadas esféricas

Como último ejemplo, sean las coordenadas cartesianas (x, y, z) en $\mathbb{R}^3$ y considere nuevas coordenadas (r, φ, θ) en $\mathbb{R}^3$ dadas por las ecuaciones

$$\begin{cases} x = r\cos\varphi\,\mathrm{sen}\,\theta, \\ y = r\,\mathrm{sen}\,\varphi\,\mathrm{sen}\,\theta \\ z = r\cos\theta \end{cases}$$

donde $r > 0$, $0 \leq \theta \leq \pi$, $0 \leq \varphi < 2\pi$ (véase la figura 5.17.)

La matriz Jacobiana para cada punto $(r, \varphi, \theta) \subset \mathbb{R}^3$ es calculada por

$$A = \begin{pmatrix} \cos\varphi\,\mathrm{sen}\,\theta & r\cos\varphi\cos\theta & -r\,\mathrm{sen}\,\varphi\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\varphi\,\mathrm{sen}\,\theta & r\,\mathrm{sen}\,\varphi\cos\theta & r\cos\varphi\,\mathrm{sen}\,\theta \\ \cos\theta & -r\,\mathrm{sen}\,\theta & 0 \end{pmatrix},$$

y su Jacobiano es

$$\frac{D(x, y, z)}{D(r, \varphi, \theta)} = det A = r^2 \mathrm{sen}\,\theta.$$

De esta manera, $det A = 0$, si y sólo sí , $r = 0$ ó $\theta = 0, \pi$, lo cual define otra vez una transformacion del espacio excepto el eje z en estas coordenadas.

Un cálculo directo nos demuestra que una transformación inversa en la región de invertibilidad, está dada por

$$\begin{cases} r = \sqrt{x^2 + y^2 + z^2} \\ \varphi = \arctan\left(\frac{y}{x}\right) \\ \theta = \arccos\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right) \end{cases}$$

Este sistema se conoce como **coordenadas co-geográficas esféricas**, y se ilustra en la figura 5.17.

EJEMPLO. Consideremos la ecuación espacial en coordenadas cartesianas dada por

$$x^2 + y^2 + z^2 - 4 = 0$$

$\triangleleft$ En coordenadas esféricas (r, φ, θ), tal ecuación se escribe

$$0 = (r\cos\varphi\ \mathrm{sen}\ \theta)^2 + (r\ \mathrm{sen}\,\varphi\ \mathrm{sen}\theta)^2 + (r\cos\theta)^2 - 4$$

$$= r^2\cos^2\varphi\ \mathrm{sen}^2\theta + r^2\ \mathrm{sen}^2\varphi\ \mathrm{sen}^2\theta + r^2\cos\theta - 4$$

$$= r^2\ \mathrm{sen}^2\theta(\cos^2\varphi + \ \mathrm{sen}^2\varphi) + r^2\cos^2\theta - 4 = r^2\ \mathrm{sen}^2\theta + r^2\cos^2\theta - 4$$

$$= r^2(\ \mathrm{sen}^2\theta + \cos^2\theta) - 4 = r^2 - 4$$

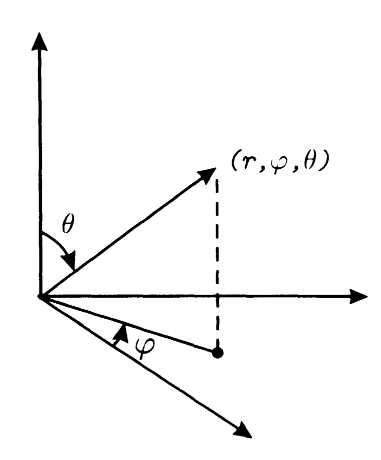


Figura 5.17: Coordenadas co-geográficas esféricas.

lo que dice que la relación inicial se lleva, bajo la transformación de coordenadas en la relación simple

$$r = 2 \quad \triangleright$$

EJEMPLO. Considérese la función

$$F(x, y, z) = \frac{1}{\ln(1 - x^2 - y^2 - z^2)}$$

definida en la región $\Omega = \{(x, y, z) | 0 < x^2 + y^2 + z^2 < 1\}$

$\triangleleft$ En coordenadas esféricas, esta transformación se reduce a

$$\widetilde{F}(r, \varphi, \theta) = F(r\cos\varphi\operatorname{sen}\theta, r\operatorname{sen}\varphi\operatorname{sen}\theta, r\cos\theta)$$

$$= \frac{1}{\ln(1 - (r\cos\varphi\operatorname{sen}\theta)^2 - (r\operatorname{sen}\varphi\operatorname{sen}\theta)^2 - (r(\cos\theta)^2)} = \frac{1}{\ln(1 - r^2)}$$

definida en la misma región que en coordenadas esféricas se escribe

$$\Omega = \{(r, \varphi, \theta) | 0 < r < 1, 0 \leq \varphi \leq 2\pi, 0 \leq \theta \leq \pi\}. \quad \triangleright$$

EJEMPLO. Consideremos la región

$$\Omega = \{(x, y, z) | 0 \leq x^2 + y^2 + z^2 \leq R^2\}$$

correspondiente a una bola sólida de radio R en el espacio $\mathbb{R}^3$.

$\triangleleft$ En coordenadas esféricas (r, φ, θ) tal región se puede escribir

$$\Omega = \{(r, \varphi, \theta) | \ 0 \leq r \leq R, \ 0 \leq \varphi \leq 2\pi, \ 0 \leq \theta \leq \pi\} \quad \triangleright$$



5.5 Campos vectoriales en $\mathbb{R}^2$ y $\mathbb{R}^3$

Considérese una región $\Omega \subset \mathbb{R}^3$, de tal forma que a cada punto $p \in \Omega$ se le asocia un vector espacial $X(p) \in \mathbb{R}^3$. Entonces diremos que en Ω está definido un campo de vectores

$$X : \Omega \to \mathbb{R}^3$$

o que en Ω actúa el **campo vectorial** X.

EJEMPLO. Sea $\Omega \subset \mathbb{R}^3$ una región que ocupa el flujo de un líquido estacionario.

$\triangleleft$ Es claro que a cada punto $p \in \Omega$ le podamos asociar su vector de velocidad $X(p)$. De esta forma, se construye el campo de vectores velocidad del flujo (**campo de velocidades**), como una función $X : \Omega \to \mathbb{R}^3$ que a cada $p \in \Omega$ le asocia el vector velocidad $X(p)$. La figura 5.18 **a.** ilustra el campo vectorial de este ejemplo. $\triangleright$

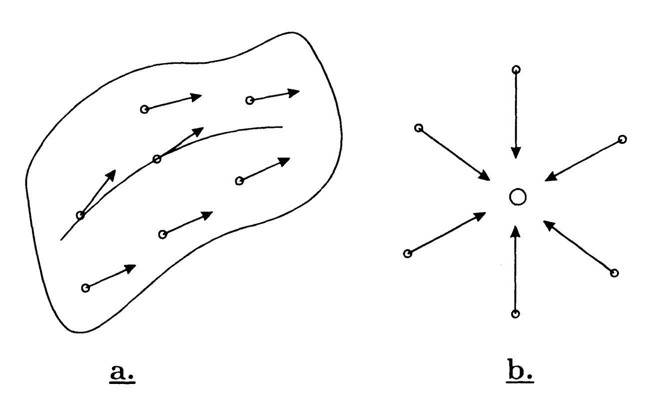


Figura 5.18: **a.** Campo de velocidades **b.** Campo gravitacional.

EJEMPLO. Consideremos la fuerza de atracción en $\mathbb{R}^3$ ejercida por un cuerpo puntual de masa m que está situado en el origen de coordenadas.

$\triangleleft$ Si $\Omega \subset \mathbb{R}^3$ es una región del espacio no conteniendo el origen, entonces para cada objeto situado en el punto $P \in \Omega$, se le asocia el vector (**fuerza gravitacional**) ejercido por su atracción hacia el cuerpo puntual de masa m.

Si se define el vector por $X(p)$, entonces se está construyendo un campo vectorial (fuerzas gravitacionales) en Ω, definido por la función $X : \Omega \to \mathbb{R}^3$ tal que a cada punto $p \in \Omega$ le asocia el vector gravitacional $X(p)$ (véase la figura 5.18 **b.** $\triangleright$



Si la asociación $p \to X(p)$ en los ejemplos anteriores es suave, como una función vectorial de argumento vectorial, diremos que los campos vectoriales son suaves.

DEFINICIÓN. Sea Ω una región del espacio $\mathbb{R}^3$. Por un **campo vectorial** suave X actuando en Ω entenderemos una función

$$X : \Omega \to \mathbb{R}^3$$

que asocia a cada punto $p \in \Omega$ el vector espacial $X(p)$, de tal forma que la asociación es suave como una función vectorial de argumento vectorial.

El la práctica, para visualizar un campo de vectores definido en una región Ω, se dibujan los vectores asociados $X(p)$ a los puntos $p \in \Omega$, teniendo a p como punto donde nace el vector $X(p)$. La figura 5.18 ilustra esta forma de visualizar los campos vectoriales de velocidades y gravitacional.

EJEMPLO. Sea un campo escalar suave arbitrario definido en una región espacial Ω,

$$F : \Omega \to \mathbb{R}$$

$\triangleleft$ Podemos asociar a cada punto $p \in \Omega$ el vector gradiente $\operatorname{grad} F(p)$, construyendo así un campo vectorial X definido por

$$X = \operatorname{grad} F : \Omega \to \mathbb{R}^3$$

donde $X(p) = \operatorname{grad} F(p)$.

Se puede comprobar fácilmente que si $\operatorname{grad} F(p) \neq (0,0,0)$, entonces es un vector ortogonal al conjunto de nivel $F^{-1}(h)$ conteniendo a p.

El campo $X = \operatorname{grad} F$ se llama el **campo gradiente** obtenido del campo escalar F. $\triangleright$

Si se considera el sistema de coordenadas x, y, z para Ω, entonces un campo de vectores X tiene la forma

$$X(x,y,z) = (P(x,y,z),\ Q(x,y,z),\ R(x,y,z)),$$

donde P, Q y R son campos escalares,

$$P, Q, R : \Omega \to \mathbb{R}$$

Por ejemplo, el campo gradiente del escalar F se escribe en las coordenadas cartesianas x, y, z,

$$X(p) = \left(\frac{\partial F}{\partial x}(p), \frac{\partial F}{\partial y}(p), \frac{\partial F}{\partial z}(p) \right)$$

siendo en este caso

$$P = \frac{\partial F}{\partial x},\ Q = \frac{\partial F}{\partial y},\ R = \frac{\partial F}{\partial z}$$

EJEMPLO. Sea el campo escalar F en el plano definido por

$$F(x,y) = x^2 + y^2$$

$\lhd$ Entonces el campo gradiente asociado al campo escalar F es

$$X(p) = \left(\frac{\partial F}{\partial x}(p), \frac{\partial F}{\partial y}(p)\right) = (2x, 2y)|_p$$

Para ilustrar cómo actúa este campo vectorial en el plano, calculamos en algunos puntos sus vectores correspondientes.

$$X(0,0) = (0,0)$$

$$X(1,0) = (2,0)$$

$$X(1,1) = (2,2)$$

$$X(0,1) = (0,2)$$

Sabemos que para $h \in \mathbb{R}$ un valor es positivo, los conjuntos de nivel del campo escalar F tienen una ecuación,

$$x^2 + y^2 = h$$

y corresponden a círculos concéntricos en el origen. Los vectores gradientes son ortogonales a ellos. La figura 5.19 ilustra este ejemplo. $\rhd$

EJEMPLO. Determinemos el campo vectorial definido por la fuerza de atracción gravitacional de un cuerpo A de masa m sobre un cuerpo B de masa 1.

$\lhd$ Supongamos que el cuerpo de masa m está localizado en $\mathbb{R}^3$. De La **ley de gravitación universal** se sigue que la fuerza X con la cual atrae el cuerpo A al cuerpo B es $X(p) = \frac{kmp}{|p|^3}$ donde $p = (x, y, z)$ es la posición del cuerpo A.

Recordemos que $|p| = \sqrt{x^2 + y^2 + z^2}$, y por lo tanto

$$X(x,y,z) = \frac{km(xe_1 + ye_2 + ze_3)}{(x^2 + y^2 + z^2)^{3/2}}$$

o equivalentemente

$$X(x,y,z) = \frac{kmxe_1}{(x^2 + y^2 + z^2)^{3/2}} + \frac{kmye_2}{(x^2 + y^2 + z^2)^{3/2}} + \frac{kmze_3}{(x^2 + y^2 + z^2)^{3/2}}$$



En este caso las funciones escalares que definen al campo vectorial son

$$P(x,y,z) = \frac{kmx}{(x^2+y^2+x^2)^{3/2}}$$

$$Q(x,y,z) = \frac{kmy}{(x^2+y^2+x^2)^{3/2}}$$

$$R(x,y,z) = \frac{kmz}{(x^2+y^2+x^2)^{3/2}} \quad \triangleright$$

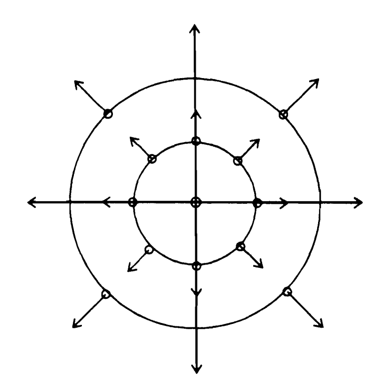


Figura 5.19: Campo gradiente de $F = x^2 + y^2$.

Mostramos que un campo vectorial crea una **"dinámica"** dentro de la región Ω contenida en $\mathbb{R}^3$ donde actúa.

EJEMPLO. Consideremos nuevamente el campo de velocidades X del líquido estacionario dentro de la región $\Omega \subset \mathbb{R}^3$. Bajo la influencia de tal campo, cualquier punto se "moverá" por la acción de la velocidad $X(p)$ del líquido, creando una trayectoria. En los puntos de tal trayectoria también se ejercen tangencialmente velocidades que son vectores del campo que actúa en Ω.

Si se conoce por $\gamma_p(t)$ la trayectoria (curva) que se realiza desde el punto p (tomado como tiempo incial $t = 0$), para tiempos $t > 0$, se tiene que la velocidad $\dot{\gamma}_p(t)$ de la curva en el punto $\gamma_p(t)$, pertenece al campo X. Esto es, el punto $\gamma_p(t)$ de la trayectoria tiene asociado al vector velocidad $X(\gamma_p(t))$ que también es la velocidad $\dot{\gamma}_p(t)$, es decir,

$$\dot{\gamma}_p(t) = X(\gamma_p(t))$$

Diremos en este caso que la curva $t \to \gamma_p(t)$, es una **curva integral** del campo vectorial X, condicionada a pasar por $\gamma_p(0) = p$.

En otras palabras, una **curva integral** del campo X de velocidades es una curva cuyos vectores tangentes pertenecen además al campo X (véase figura 5.20).

DEFINICIÓN. Sea X un campo vectorial suave en $\Omega \subset \mathbb{R}^3$, y sea

$$\gamma : (-\epsilon, \epsilon) \to \Omega, \quad \epsilon > 0$$

una curva suave definida en el intervalo $(-\epsilon, \epsilon)$ a Ω. Decimos que γ es una **curva integral** del campo X, si para todo $t \in (-\epsilon, \epsilon)$ se satisface la igualdad

$$\dot{\gamma}(t) = X(\gamma(t))$$

Esto es, γ es una curva integral de X si sus vectores tangentes $\dot{\gamma}(t)$ forman parte del campo (véase la figura 5.20).

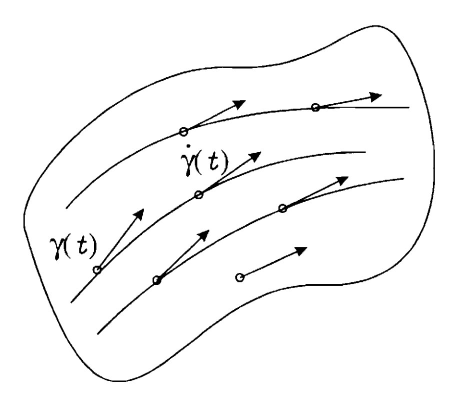


Figura 5.20: Curvas integrales de un campo vectorial

Sean x, y, z las coordenadas de Ω donde el campo X se escribe

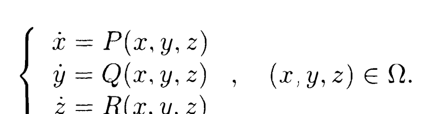
$$X(x, y, z) = (P(x, y, z), Q(x, y, z), R(x, y, z))$$

entonces la curva $\gamma(t) = (x(t), y(t), z(t))$ será una curva integral del campo X si para $t \in (-\epsilon, \epsilon)$ se cumple la cadena de igualdades

$$(\dot{x}(t), \dot{y}(t), \dot{z}(t)) = \dot{\gamma}(t) = X(\gamma(t)) = X(x(t), y(t), z(t))$$

$$= (P(x(t), y(t), z(t)), Q(x(t), y(t), z(t)), R(x(t), y(t), z(t)))$$

Esto es equivalente a que, suprimiendo la variable t (sabiendo que las coordenadas dependen de tal variable), se cumpla el siguiente sistema de ecuaciones diferenciales



$$\begin{cases} \dot{x} = P(x,y,z) \\ \dot{y} = Q(x,y,z) \\ \dot{z} = R(x,y,z) \end{cases}, \quad (x,y,z) \in \Omega.$$

Consecuentemente, para que una curva $\gamma(t)$ sea una curva integral del campo vectorial $X = (P, Q, R)$, sus coordenadas $x(t), y(t), z(t)$ deben satisfacer el sistema de ecuaciones diferenciales mencionado.

Surge el problema de la existencia de curvas integrales de un campo vectorial suave X definido en una región $\Omega \subset \mathbb{R}^3$.

El conjunto de tales curvas integrales conformarían el **flujo** asociado al campo de vectores en el cual un objeto se podría mover libremente bajo la acción de tal campo.

El siguiente teorema, el cual no demostraremos, nos garantiza la existencia de una curva integral del campo, la cual pasa por un punto condicionado y que está determinada de manera única.

TEOREMA 5.7 *Sea Ω una región de $\mathbb{R}^3$ y sea X un campo de vectores suave definido en Ω. Entonces, para el punto $p \in \Omega$*

a. *Existe una curva integral suave*

$$\gamma : (a,b) \to \Omega$$

del campo integral X, definida en un intervalo máximo $(a,b) \subset \mathbb{R}$ que contiene a 0 y tal que $\gamma(0) = p$.

b. *La curva γ es única en el sentido de que cualquier otra curva integral suave del campo*

$$\tilde{\gamma} : (a',b') \to \Omega$$

pasando por p, coincide con γ en $(a,b) \cap (a',b')$, suponiendo de inicio que $0 \in (a',b')$.

La prueba de este teorema es objeto de La Teoría de las Ecuaciones Diferenciales y sale del alcance de este libro.

Otra forma de determinar las curvas del flujo del campo vectorial

$$X = P(x,y,z)e_1 + Q(x,y,z)e_2 + R(x,y,z)e_3$$

es resolver las ecuaciones diferenciales equivalentes,

$$\frac{dx}{P(x,y,z)} = \frac{dy}{Q(x,y,z)} = \frac{dz}{R(x,y,z)}$$

pero en este caso, las soluciones están determinadas sin el parámetro temporal t.

A continuación, damos ejemplos de curvas integrales para campos vectoriales en el plano. La búsqueda de tales curvas integrales obedece en general a la solución del sistema de ecuaciones diferenciales mencionado. Se hace mención al lector que en general no es simple resolver un sistema de este tipo para encontrar las curvas integrales del campo que actúa en la región Ω, además de que los métodos de solución que se conocen merecen un trabajo aparte.

EJEMPLO. Consideremos en $\Omega = \mathbb{R}^2$ el campo de vectores constante

$$X(x, y) = (1, 0)$$

esto es, $P(x, y) = 1$ y $Q(x, y) = 0$.

$\triangleleft$ Entonces sus curvas integrales deberán de satisfacer el sistema de ecuaciones diferenciales

$$\begin{cases} \dot{x} = P = 1 \\ \dot{y} = Q = 0 \end{cases}$$

cuyas soluciones para cualquier punto $(x_0, y_0) \in \Omega$ tienen la forma

$$\gamma(t) : \begin{cases} x(t) = t + x_0 \\ y(t) = y_0 \end{cases}$$

Esta curva integral

$$t \to \gamma(t) = (x(t), y(t)) = (t + x_0, y_0)$$

corresponde a una recta paralela al eje x, que en el tiempo $t = 0$ pasa por el punto $(x_0, y_0) \in \Omega$.

La figura 5.21 ilustra el campo vectorial de este ejemplo, junto con sus curvas integrales (flujo del campo). $\triangleright$

EJEMPLO. Sea en $\Omega = \mathbb{R}^2$ el campo vectorial dado por

$$X(x, y) = (-y, x)$$

es decir, $P(x, y) = -y$, $Q(x, y) = x$.

$\triangleleft$ El sistema de ecuaciones diferenciales asociado al campo es

$$\begin{cases} \dot{x} = P(x, y) = -y \\ \dot{y} = Q(x, y) = x \end{cases}$$

Mediante métodos del Álgebra lineal se puede comprobar que la curva integral al tiempo $t = 0$ pasando por el punto $(x_0, y_0) \in \Omega$, tiene la forma

$$\gamma(t) : \begin{cases} x(t) = x_0 \cos t - y_0 \operatorname{sen} t \\ y(t) = x_0 \operatorname{sen} t + y_0 \cos t \end{cases}, \quad t \in \mathbb{R}$$

$X=(1,0)$

(x_0,y_0) (x_0+t,y_0)

Figura 5.21: Campo constante $X = (1, 0)$.

Por ejemplo, para el punto $(x_0, y_0) = (1, 1)$ se tiene que la curva integral pasando por él está definida por

$$\gamma(t) = (\cos t - \text{sen}t, \text{sen}t + \cos t)$$

Ya que

$$||\gamma(t)||^2 = (\cos t - \text{sen}t)^2 + (\text{sen}t + \cos t)^2 = 2$$

se tiene que tal curva integral por $(1, 1)$ es una circunferencia de radio $\sqrt{2}$ con centro en el origen. De manera análoga, la curva integral de X que pasa por el punto arbitrario (x_0, y_0) satisface que

$$||\gamma(t)||^2 = (x_0 \cos t - y_0 \,\text{sen}t)^2 + (x_0 \,\text{sen}t + y_0 \cos t)^2 = x_0^2 + y_0^2$$

lo que nos indica que también es una circunferencia con centro en el origen de radio $\sqrt{x_0^2 + y_0^2}$.

La figura 5.22 nos ilustra este campo vectorial junto con sus curvas integrales, que conforman una familia de círculos concéntricos, sobre los cuales se mueve el campo en dirección contraria a las manecillas del reloj. $\triangleright$

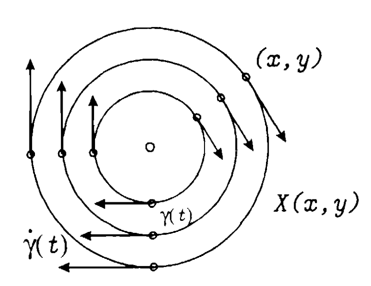


Figura 5.22: Curvas integrales del campo $X(x, y) = (-y, x)$.



Una forma más simple de obtener las curvas integrales sin la expresión explícita del parámetro t para el campo del ejemplo anterior, se muestra a continuación

EJEMPLO. Determine las líneas de flujo del campo vectorial plano del ejemplo anterior.

$$X = -ye_1 + xe_2$$

◁ Para encontrar las líneas de flujo del campo vectorial plano

$$X = P(x, y)e_1 + Q(x, y)e_2$$

debemos resolver la ecuación diferencial

$$\frac{dx}{P(x, y)} = \frac{dy}{Q(x, y)}$$

es decir,

$$\frac{dy}{dx} = \frac{Q(x, y)}{P(x, y)}$$

En nuestro ejemplo $P(x, y) = -y$, $Q(x, y) = x$, y, por lo tanto, obtenemos

$$\frac{dy}{dx} = \frac{x}{-y}$$

o equivalentemente

$$\int (-y)dy = \int xdz \iff \frac{-y^2}{2} + c_1 = \frac{x^2}{2} + c_2$$

Reacomodando los términos, concluimos que las líneas de flujo son de la forma

$$x^2 + y^2 = c$$

es decir, son circunferencias con centro en el origen. ▷

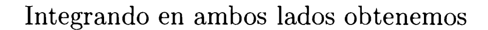
EJEMPLO. Encuentre las líneas de flujo del campo vectorial

$$X = e_1 + x(y - 1)e_2$$

◁ Las líneas de flujo se obtienen resolviendo la ecuación diferencial

$$\frac{dx}{1} = \frac{dy}{x(y - 1)}$$

es decir $xdx = \frac{dy}{y-1}$.



Integrando en ambos lados obtenemos

$$\int x dx = \int \frac{dy}{y-1},$$

es decir, siendo c una constante de integración, la solución se escribe implícitamente,

$$\frac{x^2}{2} = \ln|y-1| + c \quad \triangleright$$

EJEMPLO. (Campo gradiente). Considere un campo escalar F, y el campo gradiente asociado en la región Ω de $\mathbb{R}^3$ en las coordenadas x, y, z.

$$X(x,y,z) = \operatorname{grad} F(x,y,z) = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right)|_{(x,y,z)}$$

$\triangleleft$ Entonces las ecuaciones diferenciales para este campo gradiente vienen dadas por

$$\begin{cases} \dot{x} = \frac{\partial F}{\partial x} \\ \dot{y} = \frac{\partial F}{\partial y} \\ \dot{z} = \frac{\partial F}{\partial z} \end{cases}$$

Como es sabido, el vector tangente de una curva integral

$$\gamma(t) = (x(t), y(t), z(t))$$

es el vector

$$\dot{\gamma}(t) = \operatorname{grad} F(x(t), y(t), z(t))$$

y éste es ortogonal al conjunto de nivel donde se encuentra el punto $\gamma(t)$.

Ya que el campo gradiente es ortogonal a los conjuntos de nivel de F, entonces el flujo (conjunto de curvas integrales) de $X = \operatorname{grad} F$ es ortogonal a los conjuntos de nivel de F.

Particularmente si $F = T$ es el campo de temperaturas, entonces un objeto situado sobre un conjunto de nivel $T^{-1}(T_0)$ tenderá a moverse de manera natural hacia la dirección del gradiente $X = \operatorname{grad} T$ para ganar calor, llevado por la acción misma del campo gradiente. $\triangleright$

DEFINICIÓN. El campo vectorial X es **conservativo** si existe una función escalar F tal que $X = \nabla F$. En tal caso, llamamos a F una función **potencial** para X.

EJEMPLO. Muestre que el campo vectorial gravitacional es conservativo.

$\triangleleft$ El campo vectorial está dado por,

$$X(x,y,z) = \frac{km(xe_1 + ye_2 + ze_3)}{(x^2+y^2+z^2)^{3/2}}$$

es decir,

$$X(x,y,z) = \frac{kmxe_1}{(x^2+y^2+z^2)^{3/2}} + \frac{kmye_2}{(x^2+y^2+z^2)^{3/2}} + \frac{kmze_3}{(x^2+y^2+z^2)^{3/2}}$$

Definiendo a la función

$$F(x,y,z) = \frac{km}{\sqrt{x^2+y^2+z^2}}$$

y calculando el gradiente obtenemos,

$$\nabla F = \frac{\partial f}{\partial x}e_1 + \frac{\partial f}{\partial y}e_2 + \frac{\partial f}{\partial z}e_3$$

$$\nabla F = \frac{kmxe_1}{(x^2+y^2+z^2)^{3/2}} + \frac{kmye_2}{(x^2+y^2+z^2)^{3/2}} + \frac{kmze_3}{(x^2+y^2+z^2)^{3/2}} = X$$

con lo cual concluimos que el campo X es conservativo ▷

EJEMPLO. Determine el campo gradiente definido por el campo escalar

$$R(x,y,z) = zxe^{x^2y}$$

◁ Ya que el campo gradiente es de la forma

$$\nabla R(x,y,z) = \frac{\partial R}{\partial x}e_1 + \frac{\partial R}{\partial y}e_2 + \frac{\partial R}{\partial z}e_3$$

calculemos primero las derivadas parciales

$$\frac{\partial R}{\partial x} = ze^{x^2y} + 2x^2yz\,e^{x^2y}$$

$$\frac{\partial R}{\partial y} = x^3ze^{x^2y}$$

$$\frac{\partial R}{\partial z} = xe^{x^2y}$$

De esta manera, el campo gradiente de $R(x,y,z)$ es

$$\nabla R(x,y,z) = (ze^{x^2y} + 2x^2yze^{x^2}y)e_1 + x^3 + x^3ze^{x^2y}e_2 + xe^{x^2y}e_3 \quad \triangleright$$

EJEMPLO. Si el campo escalar

$$T(x,y,z) = \frac{23}{(1+x^2+y^2+4z^2)}$$



da la temperatura en el punto (x, y, z), determine el campo gradiente ∇T y la dirección en la cual aumenta más rápido la temperatura en el punto $(1, 1, 0)$.

$\triangleleft$ Calculemos primero las derivadas parciales de $T(x, y, z)$

$$\frac{\partial T}{\partial x} = \frac{-46x}{(1 + x^2 + y^2 + 4z^2)^2}$$

$$\frac{\partial T}{\partial y} = \frac{-46y}{(1 + x^2 + y^2 + 4z^2)^2}$$

$$\frac{\partial T}{\partial z} = \frac{-184z}{(1 + x^2 + y^2 + 4z^2)^2}$$

El campo gradiente de T es

$$\nabla T = \frac{\partial T}{\partial x} e_1 + \frac{\partial T}{\partial y} e_2 + \frac{\partial T}{\partial z} e_3$$

es decir,

$$\nabla T = \frac{-46xe_1}{(1 + x^2 + y^2 + 4z^2)^2} + \frac{-46ye_2}{(1 + x^2 + y^2 + 4z^2)^2} + \frac{184ze_3}{(1 + x^2 + y^2 + 4z^2)^2}$$

Calculemos ahora el campo gradiente en el punto $(1, 1, 0)$

$$\nabla T(1, 1, 0) = \frac{-46e_1}{9} + \frac{-46e_2}{9} + 0e_3$$

$$= -\frac{46}{9} e_1 - \frac{46}{9} e_2 + 0e_3$$

De este último se sigue que la temperatura aumenta más rápido en la dirección $\xi = \left(-\frac{46}{9}, -\frac{46}{9}, 0\right)$ $\triangleright$

El teorema 5.7 muestra que cuando en una región Ω actúa un campo de vectores, la acción de éste en cualquier punto conlleva a movimientos naturales vía las curvas integrales del campo. Cualquier otra trayectoria que no sea dentro de una curva integral deberá requerir de un "**trabajo**".

5.6 Divergencia, gradiente y rotor

En esta sección damos algunas fórmulas que son importantes para la teoría del análisis vectorial, y se hace con el espíritu de desarrollar un cálculo operacional en los campos escalares y vectoriales mediante el vector simbólico **de Hamilton**.

El **vector de Hamilton** definido por

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

actúa en los campos vectoriales X por las ecuaciones convencionales siguientes.

Para un campo vectorial $X = (P, Q, R)$ se define su **divergencia**, definida por $\operatorname{div} X$, como el campo escalar definido por

$$\operatorname{div} X = < \nabla, X >$$

donde el producto escalar se calcula convencionalmente,

$$\operatorname{div} X = < \nabla, X > = \left\langle \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right), (P, Q, R) \right\rangle = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$$

Este campo escalar $\operatorname{div} X$ mide, cuando se evalúa en el punto p, los posibles desagües o manantiales del campo X en tal punto, como se muestra en la figura 5.23. Un análisis más profundo de este concepto se realiza mediante los métodos del análisis vectorial.

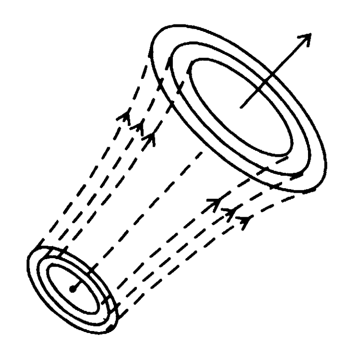


Figura 5.23: Divergencia de un campo vectorial.

Para el campo vectorial $X = (P, Q, R)$ se define su **campo de rotación** o **rotor**, definido por $\operatorname{rot} X$, como el campo vectorial

$$\operatorname{rot} X = [\nabla, X]$$

donde el producto vectorial indicado se calcula convencionalmente

$$\operatorname{rot} X = [\nabla, X] = \begin{vmatrix} e_1 & e_2 & e_3 \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ P & Q & R \end{vmatrix}$$

$$= \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) e_1 - \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right) e_2 + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) e_3$$

$$= \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) e_1 + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) e_2 + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) e_3$$

$$= \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)$$

El campo vectorial $\operatorname{rot} X$ mide, cuando se evalúa en el punto p, las posibles rotaciones que tiene el campo durante su acción en tal punto. La figura 5.24 ilustra esta idea, que puede ser estudiada más profundamente con la teoría del análisis vectorial.

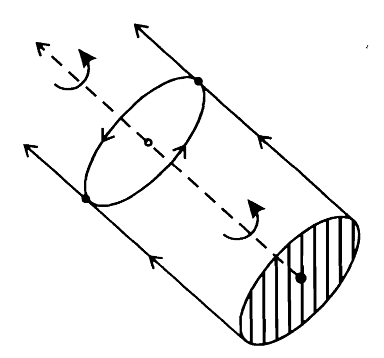


Figura 5.24: Rotor de un campo vectorial.

Si F es un campo escalar suave, podemos efectuar la operación de ∇ en F definiendo el **campo vectorial gradiente**

$$\operatorname{grad} F = \nabla F$$

donde el vector simbólico de Hamilton ∇ actúa en F mediante la igualdad

$$\operatorname{grad} F = \nabla F = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) F = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right)$$

EJEMPLO. Calcular la divergencia y el rotor del campo

$$X = (x^2, xy, z^2y) = x^2 e_1 + xy e_2 + z^2 y e_3$$

◁ Por cálculos directos se obtiene

$$\operatorname{div} X(x,y,z) = < \nabla, X > = \frac{\partial x^2}{\partial x} + \frac{\partial (xy)}{\partial y} + \frac{\partial (z^2 y)}{\partial z}$$



$$= 2x + x + 2zy = 3x + 2zy$$

$$\operatorname{rot} X = (x,y,z) = [\nabla, X] = \begin{vmatrix} e_1 & e_2 & e_3 \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ x^2 & xy & z^2y \end{vmatrix}$$

$$= \left(\frac{\partial(z^2y)}{\partial y} - \frac{\partial(xy)}{\partial z}\right) e_1 - \left(\frac{\partial(z^2y)}{\partial x} - \frac{\partial x^2}{\partial z}\right) e_2 + \left(\frac{\partial(xy)}{\partial x} - \frac{\partial x^2}{\partial y}\right) e_3$$

$$= (z^2 - 0)e_1 - (0-0)e_2 + (y-0)e_3 = (z^2, 0, y) \quad \triangleright$$

EJEMPLO. Calcular grad (div X) y div(rot X) para el campo

$$X(x,y,z) = (z^2x, x^2 - y^2, y^2 + z^2)$$

$\triangleleft$ Mediante cálculos directos se tiene

$$\operatorname{div} X(x,y,z) =< \nabla, X >= \frac{\partial(z^2x)}{\partial x} + \frac{\partial(x^2 - y^2)}{\partial y} + \frac{\partial(y^2 + z^2)}{\partial z}$$

$$= z^2 - 2y + 2z$$

$$\operatorname{rot} X = (x,y,z) = [\nabla, X] = \begin{vmatrix} e_1 & e_2 & e_3 \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ z^2x & x^2 - y^2 & y^2 + z^2 \end{vmatrix}$$

$$= \left(\frac{\partial(y^2+z^2)}{\partial y} - \frac{\partial(x^2-y^2)}{\partial z}\right) e_1 - \left(\frac{\partial(y^2+z^2)}{\partial x} - \frac{\partial(z^2x)}{\partial z}\right) e_2$$

$$+ \left(\frac{\partial(x^2-y^2)}{\partial x} - \frac{\partial(z^2x)}{\partial y}\right) e_3$$

$$= (2y - 0)e_1 - (0 - 2zx)e_2 + (2x - 0)e_3 = (2y, 2zx, 2x)$$

Consecuentemente,

$$\operatorname{grad}(\operatorname{div} X) = \nabla(z^2 - 2y + 2z)$$

$$= \left(\frac{\partial(z^2 - 2y + 2z)}{\partial x} \frac{\partial(z^2 - 2y + 2z)}{\partial y}, \frac{\partial(z^2 - 2y + 2z)}{\partial z}\right) = (0, -2, 2z + 2)$$

$$\operatorname{div}(\operatorname{rot} X) =< \nabla, (2y, 2zx, 2x) >= \frac{\partial(2y)}{\partial y} + \frac{\partial(2zx)}{2y} + \frac{\partial(2x)}{\partial z} = 0 \quad \triangleright$$

Del ejemplo anterior, para el campo dado X se obtuvo la igualdad div (rot X) $= 0$. La siguiente proposición nos indica que esta igualdad se cumple en general para cualquier campo vectorial. Se obtienen más relaciones entre los operadores grad, div y rot.

LEMA 5.1 *Para los campos vectoriales X y Y, y los campos escalares suaves F y G actuando en una región $V \subset \mathbb{R}^3$ se cumplen las identidades básicas del cálculo operacional vectorial*

a. $\mathrm{rot}\,(\mathrm{grad}\, F) \equiv 0$

b. $\mathrm{div}\,(\mathrm{rot} X) \equiv 0$

c. $\mathrm{div}\,(FX) = < \nabla F, X > + F < \nabla, X >$

d. $\mathrm{grad}\,(FG) = F \mathrm{grad}\, G + G\, \mathrm{grad}\, F$

e. $\mathrm{rot}\,(FX) = F\, \mathrm{rot}\, X + [\nabla F, X]$

f. $\mathrm{div}\,[X, Y] = < Y, \mathrm{rot}\, X > - < X, \mathrm{rot}\, Y >$

g.

$$\mathrm{div}\,(\mathrm{grad}\, F) = \Delta F$$

donde se define el operador Laplaciano

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

actuando en el campo escalar F mediante la igualdad

$$\Delta F = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + \frac{\partial^2 F}{\partial z^2}.$$

$\lhd$ Damos la demostración de algunas de las afirmaciones, dejando el resto para el lector. Como ya se mencionó, el espíritu de la prueba es puramente operacional.

a. La siguiente relación es válida,

$$\mathrm{rot}\,(\mathrm{grad}\, F) = [\nabla, \nabla F] = 0$$

pues los vectores ∇ y ∇F son paralelos.

b. Por un cálculo directo,

$$\mathrm{div}\,(\mathrm{rot}\, X) = < \nabla, [\nabla, X] > = (\nabla, \nabla, X) = 0$$

donde (, ,) es el triple producto escalar, y en la expresión dada se repite el vector ∇.

c. De la relación de divergencia, se tiene,

$$\mathrm{div}\,(FX) = < \nabla, FX > = \frac{\partial}{\partial x}(FP) + \frac{\partial}{\partial y}(FQ) + \frac{\partial}{\partial z}(FR)$$

$$= F\frac{\partial P}{\partial x} + F\frac{\partial Q}{\partial y} + F\frac{\partial R}{\partial z} + P\frac{\partial F}{\partial x} + Q\frac{\partial F}{\partial y} + R\frac{\partial F}{\partial z}$$



$$= F < \nabla, X > + < X, \nabla F >$$

f. De la definición del triple producto escalar, se tiene,

$$\mathrm{div}[X,Y] = (\nabla, X, Y) = \begin{vmatrix} \partial/\partial x & \partial/\partial y & \partial/\partial z \\ P & Q & R \\ P' & Q' & R' \end{vmatrix}$$

donde $X = (P, Q, R)$, $Y = (P', Q', R')$. Pero un cálculo directo en los determinantes (Identidad de Jacobi) muestra que

$$\begin{vmatrix} \partial/\partial x & \partial/\partial y & \partial/\partial z \\ P & Q & R \\ P' & Q' & R' \end{vmatrix} =$$

$$\begin{vmatrix} P' & Q' & R' \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ P & Q & R \end{vmatrix} - \begin{vmatrix} P & Q & R \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ P' & Q' & R' \end{vmatrix}$$

o bien

$$(\nabla, X, Y) = (X, \nabla, Y) - (Y, \nabla, X)$$

lo que prueba la afirmación **f.**

Los incisos **d.**, **e.** y **g.** se dejan al lector como ejercicio. ▷

Consideremos un cambio de coordenadas en la región $\Omega \subset \mathbb{R}^3$, dado por el sistema

$$\begin{cases} x = x(u, v, w) \\ y = y(u, v, w) \\ z = z(u, v, w), \end{cases}$$

entonces, en cada punto $p \in \Omega$, la base canónica $\{e_1, e_2, e_3\}$ se transforma, bajo La matriz Jacobiana

$$A = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix}$$

en el conjunto de vectores básicos $\{Ae_1, Ae_2, Ae_3\}$, es decir, en la base $\{f_u, f_v, f_w\}$ dada por

$$f_u = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right) = Ae_1$$

$$f_v = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}\right) = Ae_2$$

$$f_w = \left(\frac{\partial x}{\partial w}, \frac{\partial y}{\partial w}, \frac{\partial z}{\partial w} \right) = Ae_3$$

Al normalizar cada uno de esos vectores se tiene una base de vectores $\{e_u, e_v, e_w\}$, donde,

$$e_u = \frac{\left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right)}{\sqrt{\left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2}} = \frac{f_u}{||f_u||}$$

$$e_v = \frac{\left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right)}{\sqrt{\left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2}} = \frac{f_v}{||f_v||}$$

$$e_w = \frac{\left(\frac{\partial x}{\partial w}, \frac{\partial y}{\partial w}, \frac{\partial z}{\partial w} \right)}{\sqrt{\left(\frac{\partial x}{\partial w}\right)^2 + \left(\frac{\partial y}{\partial w}\right)^2 + \left(\frac{\partial z}{\partial w}\right)^2}} = \frac{f_w}{||f_w||}$$

Calculemos las normas de los vectores imágenes iniciales, por

$$h_u = \sqrt{\left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2} = ||f_u||$$

$$h_v = ||f_v||$$

$$h_w = ||f_w||$$

esto es,

$$h_u e_u = f_u, \; h_v e_v = f_v, \; h_w e_w = f_w$$

Supongamos que el conjunto de vectores $\{e_u, e_v, e_w\}$ es ortogonal, es decir, ortogonal entre sí por pares.

El elemento de desplazamiento infinitesimal vectorial $d\vec{r}$, en las coordenadas (x, y, z) está dado por

$$d\vec{r} = (dx, dy, dz) = dx e_1 + dy e_2 + dz e_3$$

Al cambiar el sistema coordenado, el **desplazamiento infinitesimal vectorial** $d\vec{r}$ en las coordenadas (u, v, w) está dado en la base $\{f_u, f_v, f_w\}$ por

$$d\vec{r} = f_u du + f_v dv + f_w dw$$

esto es, en la base unitaria $\{e_u, e_v, e_w\}$ las coordenadas de $d\vec{r}$ son ,

$$d\vec{r} = h_u e_u du + h_v e_v \, dv + h_w e_w \, dw = h_u du \, e_u + h_v dv \, e_v + h_w dw \, e_w$$



$$= (h_u du, h_v dv, h_w dw)$$

En otras palabras, bajo el cambio de coordenadas, los elementos infinitesimales son, $h_u du$ para la primer coordenada, $h_v dv$ para la segunda, y $h_w dw$ para la tercera.

Sea F un campo escalar tal que su gradiente en coordenadas (x, y, z) se escribe

$$\operatorname{grad} F = \frac{\partial F}{\partial x} e_1 + \frac{\partial F}{\partial y} e_2 + \frac{\partial F}{\partial z} e_3$$

entonces, para las coordenadas (u, v, w) se tiene la relación

$$dF = \frac{\partial F}{\partial u} du + \frac{\partial F}{\partial v} dv + \frac{\partial F}{\partial w} dw$$

$$= \left(\frac{1}{h_u}\frac{\partial F}{\partial u}\right) h_u du + \left(\frac{1}{h_v}\frac{\partial F}{\partial v}\right) h_v dv + \left(\frac{1}{h_w}\frac{\partial F}{\partial w}\right) h_w dw$$

lo que indica que las coordenadas del gradiente en la base $\{e_u, e_v, e_w\}$ son necesariamente,

$$\operatorname{grad} F = \left(\frac{1}{h_u}\frac{\partial F}{\partial u}, \frac{1}{h_v}\frac{\partial F}{\partial v}, \frac{1}{h_w}\frac{\partial F}{\partial w}\right)$$

Cálculos más complicados nos demuestran que para un campo vectorial $X = (P, Q, R)$, el cambio de coordenadas nos lleva la divergencia de X en las coordenadas x, y, z,

$$\operatorname{div} X = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$$

en la expresión

$$\operatorname{div} X = \frac{1}{h_u h_v h_w} \left[\frac{\partial}{\partial u} (P\, h_v h_w) + \frac{\partial}{\partial v} (Q h_u\, h_w) + \frac{\partial}{\partial w} (R\, h_u h_v)\right]$$

en las coordenadas (u, v, w).

Por otro lado, el rotor del campo X en las coordenadas x, y, z dado por

$$\operatorname{rot} X = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) e_1 + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) e_2 + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) e_3$$

se transforma en

$$\operatorname{rot} X = \frac{1}{h_v h_w} \left(\frac{\partial(R\, h_w)}{\partial v} - \frac{\partial(Q\, h_v)}{\partial w}\right) e_u$$
$$+ \frac{1}{h_w h_u} \left(\frac{\partial(P\, h_u)}{\partial w} - \frac{\partial(R\, h_w)}{\partial u}\right) e_v + \frac{1}{h_u h_v} \left(\frac{\partial(Q\, h_v)}{\partial x} - \frac{\partial(P\, h_u)}{\partial z}\right) e_w$$



en las coordenadas (u, v, w).

EJEMPLO. Sean las coordenadas esféricas co-geográficas

$$\begin{cases} x = r\cos\varphi \operatorname{sen}\theta \\ y = r\operatorname{sen}\varphi\operatorname{sen}\theta, \\ z = r\cos\theta \end{cases}$$

$\lhd$ Ya se ha demostrado que

$$A = \begin{pmatrix} \cos\varphi\operatorname{sen}\theta & -r\operatorname{sen}\varphi\operatorname{sen}\theta & r\cos\varphi\cos\theta \\ \operatorname{sen}\varphi\operatorname{sen}\theta & r\cos\varphi\operatorname{sen}\theta & r\operatorname{sen}\varphi\cos\theta \\ \cos\theta & 0 & -r\operatorname{sen}\theta \end{pmatrix}$$

lo cual implica, haciendo $u = r$, $v = \varphi$, $w = \theta$,

$$f_r = (\cos\varphi\operatorname{sen}\theta,\ \operatorname{sen}\varphi\operatorname{sen}\theta,\ \cos\theta)$$

$$f_\varphi = (-r\operatorname{sen}\varphi\operatorname{sen}\theta,\ r\cos\varphi\operatorname{sen}\theta,\ 0)$$

$$f_\theta = (r\cos\varphi\cos\theta,\ r\operatorname{sen}\varphi\cos\theta,\ -r\operatorname{sen}\theta)$$

Al calcular las normas se tiene

$$h_r = ||f_r|| = 1$$

$$h_\varphi = ||f_\varphi|| = r\operatorname{sen}\theta$$

$$h_\theta = ||f_\theta|| = r$$

De esta manera, los vectores básicos en las coordenadas (r, φ, θ) son

$$e_r = \frac{f_r}{h_r} = (\cos\varphi\operatorname{sen}\theta,\ \operatorname{sen}\varphi\operatorname{sen}\theta, \cos\theta)$$

$$e_\varphi = \frac{f_\varphi}{h_r} = (-\operatorname{sen}\varphi, \cos\varphi,\ 0)$$

$$e_\theta = (\cos\theta,\ \operatorname{sen}\varphi\cos\theta,\ -\operatorname{sen}\theta)$$

y un cálculo directo prueba que $\{e_r, e_\varphi, e_\theta\}$ es ortogonal.

De esta forma, para un campo escalar F, en las coordenadas (r, φ, θ) se tiene que

$$\operatorname{grad} F = \left(\frac{1}{h_r}\frac{\partial F}{\partial r}, \frac{1}{h_\varphi}\frac{\partial F}{\partial \varphi}, \frac{1}{h_\theta}\frac{\partial F}{\partial \theta}\right)$$

$$= \left(\frac{\partial F}{\partial r}, \frac{1}{r\operatorname{sen}\theta}\frac{\partial F}{\partial \varphi}, \frac{1}{r}\frac{\partial F}{\partial \theta}\right)$$

Mientras que para un campo vectorial $X = (P, Q, R)$ se cumplen las fórmulas

$$\operatorname{div} X = \frac{1}{h_r h_\varphi h_\theta}\left[\frac{\partial(P h_\varphi h_\theta)}{\partial r} + \frac{\partial(Q h_r h_\theta)}{\partial \varphi} + \frac{\partial(R h_r h_\varphi)}{\partial \theta}\right]$$

$$= \frac{1}{r^2 \operatorname{sen}\theta}\left[\frac{\partial(r^2 \operatorname{sen}\theta P)}{\partial r} + \frac{\partial(rQ)}{\partial \varphi} + \frac{\partial(r \operatorname{sen}\theta R)}{\partial \theta}\right]$$

$$= \frac{1}{r^2 \operatorname{sen}\theta}\left[\operatorname{sen}\theta \frac{\partial(r^2 P)}{\partial r} + r\frac{\partial Q}{\partial \varphi} + r\frac{\partial(\operatorname{sen}\theta R)}{\partial \theta}\right]$$

$$= \frac{1}{r^2}\frac{\partial(r^2 P)}{\partial r} + \frac{1}{r \operatorname{sen}\theta}\frac{\partial Q}{\partial \varphi} + \frac{1}{r \operatorname{sen}\theta}\frac{\partial(\operatorname{sen}\theta R)}{\partial \theta}$$

$$\operatorname{rot} X = \frac{1}{h_\varphi h_\theta}\left(\frac{\partial(R h_\theta)}{\partial \varphi} - \frac{\partial(Q h_\varphi)}{\partial \theta}\right) e_r$$

$$+\frac{1}{h_\theta h_r}\left(\frac{\partial(P h_r)}{\partial \theta} - \frac{\partial(R h_\theta)}{\partial r}\right) e_\varphi + \frac{1}{h_r h_\varphi}\left(\frac{\partial(Q h_\varphi)}{\partial r} - \frac{\partial(P h_r)}{\partial_\varphi}\right) e_\theta$$

$$= \frac{1}{r^2 \operatorname{sen}\theta}\left(\frac{\partial(rR)}{\partial \varphi} - \frac{\partial(r \operatorname{sen}\theta Q)}{\partial \theta}\right) e_r$$

$$+\frac{1}{r}\left(\frac{\partial P}{\partial \theta} - \frac{\partial(rR)}{\partial r}\right) e_\varphi + \frac{1}{r \operatorname{sen}\theta}\left(\frac{\partial(r \operatorname{sen}\theta Q)}{\partial r} - \frac{\partial P}{\partial \varphi}\right) e_\theta$$

$$= \frac{1}{r^2 \operatorname{sen}\theta}\left(r\frac{\partial R}{\partial \varphi} - r\frac{\partial(\operatorname{sen}\theta R)}{\partial \theta}\right) e_r$$

$$+\frac{1}{r}\left(\frac{\partial P}{\partial \theta} - \frac{\partial(rR)}{\partial r}\right) e_\varphi + \frac{1}{r \operatorname{sen}\theta}\left(\operatorname{sen}\theta\frac{\partial(rQ)}{\partial r} - \frac{\partial P}{\partial \varphi}\right) e_\theta$$

Dejamos como ejercicio al lector verificar que en las coordenadas (u, v, w) se cumple la siguiente igualdad para el Laplaciano de un campo escalar F.

$$\Delta F = \operatorname{div} \operatorname{grad} F$$

$$= \frac{1}{h_u h_v h_w}\left[\frac{\partial}{\partial u}\left(\frac{h_v h_w}{h_u}\frac{\partial F}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{h_u h_w}{h_v}\frac{\partial F}{\partial v}\right) + \frac{\partial}{\partial w}\left(\frac{h_u h_v}{h_w}\frac{\partial F}{\partial w}\right)\right]$$

Consecuentemente, para las coordenadas esféricas se tiene, para un campo escalar F,

$$\Delta F = \frac{1}{r^2 \operatorname{sen}\theta}\left[\frac{\partial}{\partial r}\left(r^2 \operatorname{sen}\theta \frac{\partial F}{\partial r}\right) + \frac{\partial}{\partial \varphi}\left(\frac{1}{\operatorname{sen}\theta}\frac{\partial F}{\partial \varphi}\right) + \frac{\partial}{\partial \theta}\left(\operatorname{sen}\theta\frac{\partial F}{\partial \theta}\right)\right]$$

$$= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial F}{\partial r}\right) + \frac{1}{r^2 \operatorname{sen}^2\theta}\frac{\partial^2 F}{\partial \varphi^2} + \frac{1}{r^2 \operatorname{sen}\theta}\frac{\partial}{\partial \theta}\left(\operatorname{sen}\theta\frac{\partial F}{\partial \theta}\right) \quad \triangleright$$



La ecuación del calor

Consideremos un cuerpo físico $\Omega \subset \mathbb{R}^3$, y designemos su temperatura en el punto (x, y, z) en el instante t por el campo escalar

$$u = u(x, y, z, t)$$

Afirmamos que el campo u satisface la ecuación diferencial parcial lineal de segundo orden

$$\frac{\partial u}{\partial t} = a^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right)$$

donde $(x, y, z) \in \Omega$ y $a \in \mathbb{R}$.

Al utilizar el operador de Laplace,

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

lo anterior equivale a que la función u debe satisfacer la ecuación

$$\frac{\partial u}{\partial t} = a^2 \Delta u$$

llamada la **ecuación de conducción de calor**.

$\lhd$ Al tomar un cubo infinitesimal σ del cuerpo Ω (véase la figura 5.25), la cantidad de calor que atraviesa la cara izquierda de σ de derecha a izquierda durante el tiempo $(t, t + \Delta t)$ es, hasta, un infinitésimo

$$\alpha \frac{\partial u}{\partial x}(x, y, z, t) \Delta y \, \Delta z \, \Delta t$$

donde α es el coeficiente de conductibilidad térmica del cuerpo que se considera constante en cualquiera de sus puntos.

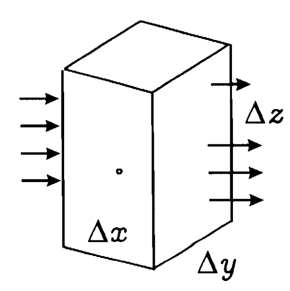


Figura 5.25: Calor atravesando el cubo infinitesimal en Ω.



Lo anterior se debe a que la cantidad indicada de calor es proporcional al número α, al área $\Delta y\, \Delta z$ de la cara que se examina, al incremento de tiempo Δt y a la velocidad de variación de la temperatura en la dirección del eje x que es igual a la derivada parcial $\frac{\partial u}{\partial x}$. La derivada parcial varía dentro de los límites de la cara, pero, despreciando infinitésimos de orden superior, se puede suponer que en toda la cara es igual a $\frac{\partial u}{\partial x}$ en el punto (x, y, z).

La cantidad de calor que pasa por la cara derecha de σ de derecha a izquierda es, análogamente, igual a

$$\alpha \frac{\partial u}{\partial x}(x + \Delta x, y, z, t)\Delta y\, \Delta z\, \Delta t$$

De esta forma, la cantidad de calor que entra en el cubo σ por sus caras izquierda y derecha durante el lapso de indicado es igual a

$$\alpha \frac{\partial u}{\partial x}(x + \Delta x, y, z, t)\Delta y\, \Delta z\, \Delta t - \alpha \frac{\partial u}{\partial x}(x, y, z)\Delta y\, \Delta z\, \Delta t$$

pero, por el teorema del valor medio para u en la variable x, se tiene que

$$\alpha \frac{\partial u}{\partial x}(x + \Delta x, y, z, t)\Delta y\, \Delta z\, \Delta t - \alpha \frac{\partial u}{\partial x}(x, y, z)\Delta y\, \Delta z\, \Delta t$$

$$= \alpha \frac{\partial^2 u}{\partial x^2}(x, y, z, t)\Delta x\, \Delta y\, \Delta z\, \Delta t.$$

En virtud de que en las otras caras del cubo ocurre igual, entonces la cantidad total de calor que entra en σ en el lapso $(t, t + \Delta t)$ es, la suma de las cantidades de calor que entra durante este tiempo por todas las caras de σ.

$$\alpha \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) \Delta x\, \Delta y\, \Delta z\, \Delta t.$$

Por las mismas consideraciones, hasta un infitésimo del volumen $\Delta x\, \Delta y\, \Delta z$ de σ, este número (cantidad de calor) es igual también a

$$\beta \frac{\partial u}{\partial t} \Delta x\, \Delta y\, \Delta z\, \Delta t$$

donde β es el calor específico del cuerpo que suponemos constante en todos sus puntos.

Igualando, obtenemos, después de las simplificaciones, la ecuación diferencial parcial lineal de segundo orden mencionada, donde

$$a^2 = \frac{\alpha}{\beta}.$$



Se ha mostrado que la temperatura del cuerpo Ω es la función $u = u(x, y, z, t)$ que satisface la ecuación de conducción del calor, donde a^2 es una constante positiva. Nos hemos limitado al caso cuando el cuerpo tiene en todos sus puntos un calor específico constante y un coeficiente de conductibilidad que no varía.

La ecuación diferencial obtenida tiene un conjunto infinito de soluciones, y para encontrar entre ellas una solución determinada es necesario imponer sobre la función u condiciones adicionales, llamadas **condiciones iniciales y de frontera**. Este problema sale del alcance de este trabajo.

La distribución del calor en un cuerpo se llama **estacionaria** si la temperatura u del cuerpo depende de la posición del punto (x, y, z) y no depende del tiempo t. Es decir, si $u = u(x, y, z)$, entonces

$$\frac{\partial u}{\partial t} = 0$$

lo que implica que el campo escalar u satisface la ecuación diferencial parcial lineal de segundo orden

$$\Delta u = 0.$$

DEFINICIÓN. La función $u(x, y, z)$ se llama **armónica** sobre la región Ω si tiene derivadas parciales continuas de segundo orden sobre Ω y satisface sobre Ω la ecuación

$$\Delta u = 0.$$

llamada la **ecuación de Laplace**.

LEMA 5.2 *Suponga que la región acotada Ω tiene como frontera suave a la superficie S sobre la cual se dá la función continua $f(x, y, z)$. Entonces en la cerradura $\bar{\Omega}$ existe una única función continua $u(x, y, z)$, armónica sobre la región Ω, tal que*

$$u|_S = f(x, y, z)$$

Este resultado tiene una interpretación física sencilla. Si sobre la frontera superficial S del cuerpo Ω se mantiene una temperatura u, tal que $u|_S = f(x, y, z)$ es la función continua dada sobre S, entonces dentro del cuerpo se establece una única temperatura bien determinada, armónica $u(x, y, z)$.

Desde el punto de vista físico, esta afirmación es obvia, aunque puede ser demostrada matemáticamente. Este problema, llamado el **problema de Dirichlet**, está parcialmente investigado y se conocen diferentes métodos numéricos de solución.



Ejercicios.

1. Sean las transformaciones lineales

$$L_1(x,y,z) = (x+2y+3z, 4x+5y+6z)$$

$$L_2(u,v,w) = (u+3v+5w, 6u+7v+9w)$$

Calcular $L_1 \circ L_2$ y $L_2 \circ L_1$, usando sus matrices asociadas en la base canónica.

2. Hacer lo mismo que el ejercicio **1.** para las transformaciones lineales

$$L_3(x,y,z) = (x+y, y+z, z+x)$$

$$L_4(u,v,w) = (v+w, u+w, u+v)$$

es decir, calcular $L_3 \circ L_4$ y $L_4 \circ L_3$.

3. Sean las transformaciones lineales,

$$L_5(x,y) = (2x-y, x+3y, x)$$

$$L_6(x,y,z) = (x+y+z, x-y-z)$$

y considérense las transformaciones L_1, L_2, L_3 y L_4 en los ejercicios **1.** y **2.**. Calcular las transformaciones lineales

a. $L_4 \circ L_5$

b. $L_5 \circ L_6$

c. $L_2 \circ L_5$

Calcular la aplicación de cada transformación lineal dada en el vector indicado.

d. $L_5(\xi)$, si $\xi = (-1,1)$

e. $L_6(\xi)$, si $\xi = (1,0,1)$

f. $L_3(\xi)$, si $\xi = (-1,2,3)$

g. $L_4 \circ L_5(\xi)$, si $\xi = (1,2)$

h. $L_5 \circ L_6(\xi)$, si $\xi = (3,-2,1)$

4. Dar los dominios de definición (regiones) de los siguientes campos vectoriales.

a. $f_1(x,y) = (x^2+y^2,\ xy-x^2)$

b. $f_2(x,y) = \left(\sqrt{x^2-y^2}, xy^3\ \ln(xy)\right)$

c. $f_3(x,y,z) = \left(xyz^2, \frac{x}{yz}\right)$



d. $f_4(x,y,z) = (\operatorname{arcsen}(x+y+z), xy,\ x^2+y^2+z^2)$

5. Utilizando las funciones del ejercicio **4.** calcular las composiciones indicadas.

a. $f_3 \circ f_2$

b. $f_1 \circ f_3$

c. $f_3 \circ f_4$

d. $f_4 \circ f_3$

e. $f_4 \circ f_2$

6. Calcular la matriz Jacobiana de cada uno de los incisos del ejercicio **4.**

7. Calcular la matriz Jacobiana de la función dada en el punto p indicado.

a. $f(x,y,z) = (x^2+y^2,\ x\cos yz)$ en $p = (1,0,\pi/2)$
b. $f(x,y) = (e^x \operatorname{sen} x,\ e^y \cos x)$ en $p = (0,0)$
c. $f(x,y,z) \left(\frac{x^2}{1+y}, 3x+z^2, y^2z\right)$ en $p = (1,1,1)$
d. $f(x,y) = (\ln xy,\ \operatorname{sen} xy, x^2+y)$ en $p = (2,1)$

8. **a.** Hallar los puntos del dominio de la función **7.** **b.** donde el Jacobiano se anula.

b. Hallar los puntos del dominio de la función **8.** **c.** donde el Jacobiano se anula.

9. Calcular, utilizando La regla de la cadena, las matrices Jacobianas de las composiciones de todos los incisos del ejercicio **5.**

10. ¿Cuáles de las siguientes funciones se pueden considerar un cambio de coordenadas local, en el punto que se indica?

a. $f(x,y) = (x^2y+1, x^2+y^2)$ en $p = (1,1)$
b. $f(x,y) = (x^2+y^2, \sqrt{y}+\sqrt{x})$ en $p = (4,16)$
c. $f(x,y) = (e^{xy}, \ln x)$ en $p = (1,\ln 2)$
d. $f(x,y,z) = (xz, xy, yz)$ en $p = (1,4)$
f. $f(x,y,z) = (y\cos x, y\operatorname{sen} x, z)$ en $p = (\pi/4, 4, 3)$

11. Dibuje e integre, sin usar el parámetro temporal, los siguientes campos vectoriales.

a. $X(x,y) = (x,y)$

b. $X(x,y) = (-x,y)$

c. $X(x,y) = (x,-y)$

d. $X(x,y) = (-x,-y)$

e. $X(x,y) = (x^2,y)$

f. $X(x,y) = (x^2, -y)$

g. $X = \operatorname{grad} F$, donde $F(x,y) = x^2 - y^2$

12. a. Calcule la divergencia y el rotor de todos los campos vectoriales del ejercicio **11.**

b. Calcule la divergencia y el rotor de los campos vectoriales siguientes.

i. $X_1(x,y,z) = (y - x,\ x + y,\ yz)$

ii. $X_2(x,y,z) = (x^2 e^{-yz},\ y^2 e^{-xz},\ z^2 e^{-xy})$

iii. $X_3(x,y,z) = (z^2 y - x,\ x^2 z + y,\ y^2 x - z)$

iv. $X_4(x,y,z) = (x^x \operatorname{sen} zy,\ e^y \cos xz,\ e^z \operatorname{sen} yx)$

13. Verifique las igualdades siguientes para los campos vectoriales del ejercicio **12.**

a. $\operatorname{div}(\operatorname{rot} X) = 0, \quad X = X_1, X_2, X_3, X_4$

b. $\operatorname{div}[X_1, X_2] = < X_2, \operatorname{rot} X_1 > - < X_1, \operatorname{rot} X_2 >$

c. $\operatorname{div}[X_3, X_4] = < X_4, \operatorname{rot} X_3 > - < X_3, \operatorname{rot} X_4 >$

14. Si $F(x,y,z) = x + y^2 + z^2$, verifique las igualdades siguientes, usando los campos vectoriales del ejercicio **12.**

a. $\operatorname{rot}(FX) = F \operatorname{rot} X + [\nabla F, X], \quad X = X_1, X_2$

b. $\operatorname{rot}(\operatorname{grad} F) = 0$

c. $\operatorname{div}(\operatorname{grad} F) = \Delta F$

15. Demostrar que en las coordenadas (u,v,w) se cumple la igualdad siguiente para el Laplaciano de un campo escalar F arbitrario.

$$\Delta F = \frac{1}{h_u h_v h_w}\left[\frac{\partial}{\partial u}\left(\frac{h_v h_w}{h_u}\frac{\partial F}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{h_u h_w}{h_v}\frac{\partial F}{\partial v}\right) + \frac{\partial}{\partial w}\left(\frac{h_u h_v}{h_w}\frac{\partial F}{\partial w}\right)\right]$$

16. Calcule la divergencia y el rotor de cualquier campo vectorial X, y el gradiente de un campo escalar, en los siguientes sistemas de coordenadas.

a. En coordenadas cilíndricas.

b. En las coordenadas (u,v,w) donde

$$\begin{cases} x = ue^{vw} \\ y = ve^{uw} \\ z = we^{uv} \end{cases}$$

c. En las **coordenadas pseudoesféricas** (r, χ, φ) donde

$$\begin{cases} x = r \cosh \chi \\ y = r \operatorname{senh} \chi \cos \varphi \\ z = r \operatorname{senh} \chi \operatorname{sen} \varphi, \end{cases}$$

donde $r > 0$, $\chi > 0$, $0 \leq \varphi \leq 2\pi$.



Apéndices


Casa abierta al tiempo



Capítulo 6

Elementos Básicos de Superficies en $\mathbb{R}^3$

En este capítulo introducimos el concepto básico de superficie en el espacio euclidiano $\mathbb{R}^3$. Damos ejemplos de las superficies más comúnmente conocidas y un tratamiento general de su estudio.

Utilizamos una notación para definir a los planos coordenados de la siguiente manera.

$\mathbb{R}^2_{x,y}$ definirá al plano (x,y) contenido en $\mathbb{R}^3$, $\mathbb{R}^2_{x,z}$, definirá al plano (x,z) contenido en $\mathbb{R}^3$, mientras que $\mathbb{R}^2_{y,z}$ definirá al plano (y,z) contenido en $\mathbb{R}^3$.

DEFINICIÓN. A los puntos $(x,y,z) \in \mathbb{R}^3$ que satisfagan una relación real del tipo

$$R(x,y,z) = 0$$

se les dirá que conforman una **superficie ordinaria** en $\mathbb{R}^3$.

EJEMPLO. (El plano). Un conjunto de puntos en el espacio que satisfacen una ecuación del tipo

$$ax + by + cz = d$$

se dice que forman un plano con vector normal $n = (a, b, c)$.

$\lhd$ Tales puntos también satisfacen la relación

$$R(x,y,z) = ax + by + cz - d = 0$$

lo que asegura que un plano es una superficie (plana) en $\mathbb{R}^3$. $\rhd$

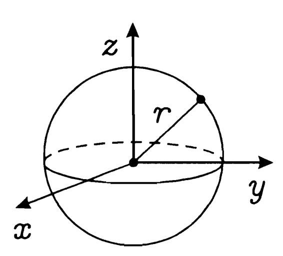


Figura 6.1: Superficie esférica en $\mathbb{R}^3$.

EJEMPLO. (**La esfera** de radio r) Sea $r > 0$ un número real dado.

$\lhd$ Los puntos $(x, y, z) \in \mathbb{R}^3$ tales que distan r del origen conforman una superficie esférica satisfaciendo la ecuación

$$\sqrt{x^2 + y^2 + z^2} = r$$

o equivalentemente,

$$x^2 + y^2 + z^2 = r^2$$

En otras palabras, los puntos sobre tal esfera, deberán satisfacer la relación:

$$R(x, y, z) = x^2 + y^2 + z^2 - r^2 = 0$$

lo que hace que la esfera con radio r y centro en el origen sea una superficie en $\mathbb{R}^3$. Esto se ilustra en la figura 6.1.

Para muestra, la ecuación

$$x^2 + y^2 + z^2 = 16$$

describe a una superficie esférica de radio 4 con centro en $(0, 0, 0)$. $\rhd$

6.1 Superficies de revolución

Consideremos en el plano x, z una curva γ que es la gráfica de una función $z = f(x)$, tal que no corta al eje z.

Al hacer girar γ en torno al eje z se obtiene una **superficie de revolución** S contenida en $\mathbb{R}^3$.

Afirmamos que S es una superficie ordinaria según nuestra definición.

$\lhd$ Sea el punto $(x_0, y_0, z_0) \in S$. Afirmamos que satisface una relación del tipo

$$R(x_0, y_0, z_0) = 0$$



Cierto es que (x_0, y_0, z_0) está en el círculo obtenido al hacer girar el punto (x_0, z_0) de la curva γ. Pero entonces,

$$z_0 = f(x_0)$$

y todos los puntos del círculo de radio x_0 están a la misma altura z_0 y contenidos en el plano $R^2_{x,y}$ en un círculo de radio x_0. Es decir, tienen una ecuación en el plano $z = z_0$ dada por $x^2 + y^2 = x_0^2$ como se muestra en la figura 6.2.

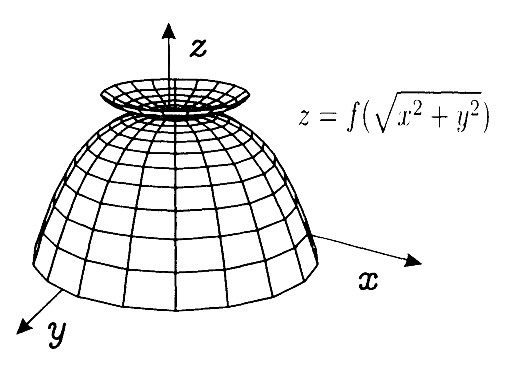


Figura 6.2: Superficie de revolución.

Esto implica que sobre tal círculo se cumple la igualdad

$$z_0 = f(\sqrt{x^2 + y^2})$$

Ya que (x_0, y_0, z_0) es arbitrario, entonces cada punto $(x, y, z) \in S$ debe satisfacer la ecuación

$$z = f(\sqrt{x^2 + y^2})$$

es decir, los puntos $(x, y, z) \in S$ satisfacen además la relación

$$R(x, y, z) = z - f(\sqrt{x^2 + y^2}) = 0$$

lo que hace de S una superficie. $\triangleright$

De igual forma, si $z = f(y)$ y se hace girar γ alrededor del eje z, se obtiene la misma relación. Esto es, si $z = f(x)$ ó $z = f(y)$ para obtener la ecuación de la superficie S obtenida al girar la gráfica γ de f alrededor del eje z es necesario reemplazar, en un caso,

$$x \to \sqrt{x^2 + y^2}$$



o bien, en el otro caso,

$$y \to \sqrt{x^2 + y^2}$$

EJEMPLO. Consideremos la curva

$$z = x^2$$

que es la ecuación de la parábola en el plano x, z con $x \geq 0$.

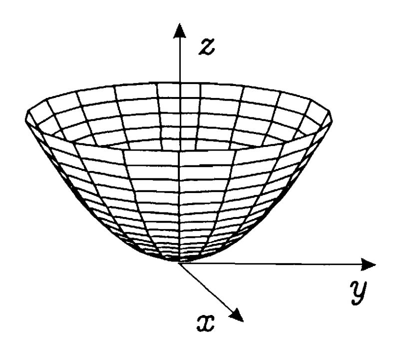


Figura 6.3: Paraboloide circular $z = x^2 + y^2$.

La ecuación de la superficie de la revolución S (**paraboloide circular**) obtenida al hacer girar tal parábola alrededor del eje z es

$$z = (\sqrt{x^2 + y^2})^2 = x^2 + y^2$$

o equivalentemente, la relación

$$x^2 + y^2 - z = 0$$

Véase la figura 6.3 que describe a esta superficie.

EJEMPLO. Sea

$$z = \frac{1}{y}$$

la hipérbola canónica en el plano y, z con $y > 0$. La ecuación de la superficie de revolución S obtenida al hacer girar alrededor del eje z es

$$z = \frac{1}{\sqrt{x^2 + y^2}}$$

que se logra al sustituir $y \to \sqrt{x^2 + y^2}$.



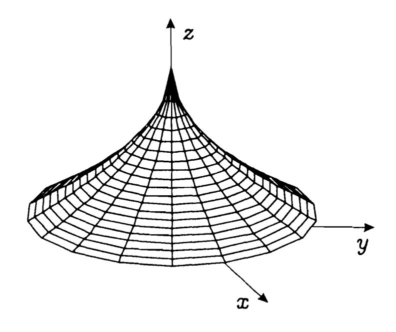


Figura 6.4: Hiperboloide circular.

Tal superficie se llama un **hiperboloide circular** y se ilustra en la figura 6.5.

EJEMPLO. Sea el semicírculo en el plano $\mathbb{R}^2_{x,y}$ dado por la ecuación

$$x = \sqrt{R^2 - y^2}$$

donde la variable $x \leq 0$.

Al hacer girar en torno al eje y (intercambiando) reemplazando en este caso la variable

$$x \rightarrow \sqrt{z^2 + x^2}$$

se obtiene la superficie de revolución

$$\sqrt{z^2 + x^2} = \sqrt{R^2 - y^2}$$

o bien,

$$x^2 + y^2 + z^2 = R^2$$

que es la ya conocida esfera de radio R con centro en $(0, 0, 0)$. $\triangleright$

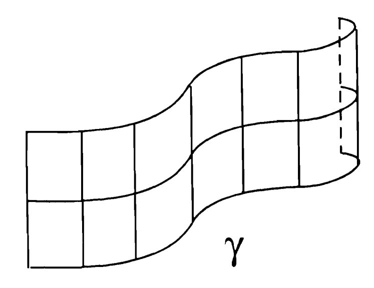


Figura 6.5: Superficie cilíndrica.



6.2 Superficies cilíndricas

Sea γ una curva contenida en un plano coordenado y sea ℓ el eje perpendicular a tal plano.

Si por cada punto $(x_0, y_0, z_0) \in \gamma$ pasamos una recta paralela a ℓ obtenemos una **superficie cilíndrica** S. La curva γ se llama **generatriz** de S y ℓ se llama la **directriz**.

Se afirma que S es superficie en nuestro contexto.

$\lhd$ Sin pérdida de generalidad, supóngase que γ está en el plano $\mathbb{R}^2_{x,y}$. Por ser curva en $\mathbb{R}^2_{x,y}$ las coordenadas x, y satisfacen una ecuación $R(x, y) = 0$

Sea $(x, y, z) \in S$ entonces $(x, y) \in \gamma$ lo que implica que $R(x, y) = 0$.

Claramente el punto (x, y, z), omitiendo a la variable z, también la satisface, es decir, si se define (abusando de la notación) a la relación

$$R(x, y, z) = R(x, y) = 0$$

entonces S es una superficie (véase la figura 6.5).

EJEMPLO. Sea la curva plana

$$x^2 + y^2 = 16$$

un círculo de radio 4 con centro en el origen. Esta genera en el espacio la superficie cilíndrica S definida por la misma ecuación, quedando libre la variable z.

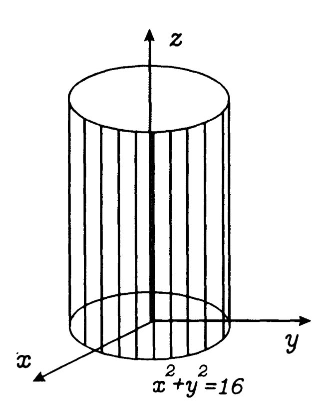


Figura 6.6: Cilindro circular recto.

Se llama **cilindro circular recto** de radio 2 con centro en el origen y con directriz al eje z (véase figura 6.6).



EJEMPLO. Sea la curva sinusoidal

$$y = \operatorname{sen} x,$$

$x \in [-\pi, \pi]$. Tal curva genera la superficie cilíndrica en la misma ecuación, siendo z libre. Se llama **cilindro sinusoidal** con directriz al eje z. La figura 6.7 ilustra a esta superficie.

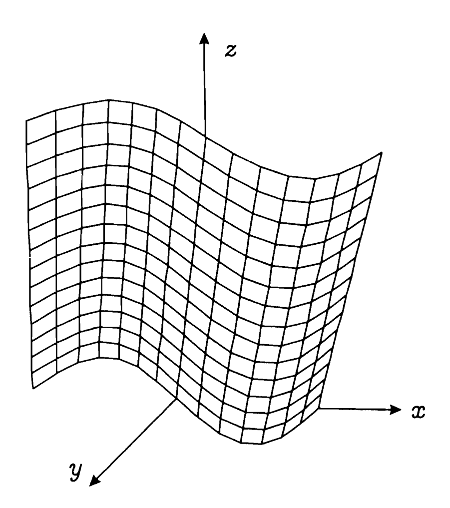


Figura 6.7: Cilindro sinusoidal.

EJEMPLO. Sea γ la gráfica de la función $z = e^y$ en el plano $\mathbb{R}^2_{z,y}$. Esta curva genera la superficie cilíndrica

$$z - e^y = 0$$

con la variable x libre.

Se llama **cilindro exponencial** con directriz el eje x (véase la figura 6.8).

6.3 Superficies cónicas

Dada una curva γ en el espacio $\mathbb{R}^3$ y un punto v, se define una superficie S al tomar el haz de rectas por v que intersectan también a γ.

Se llama una **superficie cónica** con **vértice** v y **base** γ. La figura 6.9 ilustra a este tipo de superficies.

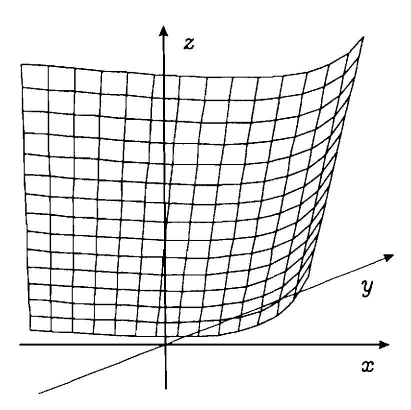


Figura 6.8: Cilindro exponencial.

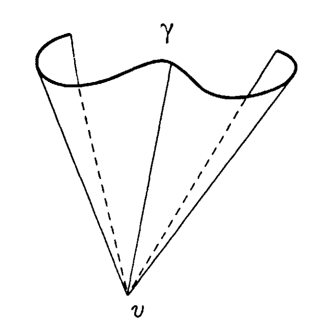


Figura 6.9: Superficie cónica.

Una relación $R(x, y, z)$ en las variables x, y, z, se dice homogénea de grado s, si para cualquier $t \in \mathbb{R}$ y cualquier punto (x, y, z) se cumple la igualdad

$$R(tx, ty, tz) = t^s R(x, y, z)$$

Para muestra, la relación

$$R(x, y, z) = x^2 + y^2 - z^2$$

es homogénea de grado 2. En efecto,

$$\triangleleft \quad R(tx, ty, tz) = (tx)^2 + (ty)^2 - (tz)^2 = t^2x^2 + t^2y^2 - t^2z^2$$

$$= t^2(x^2 + y^2 - z^2) = t^2R(x, y, z) \quad \triangleright$$



Otro ejemplo, es la relación

$$R(x,y,z) = xy^2 + x^2y + y^3 - z^2y$$

la cual es una relación homogénea de grado 3, como se muestra en el siguiente cálculo.

$$\triangleleft \quad R(tx,ty,tz) = tx(ty)^2 + (tx)^2(ty) + (ty)^3 - (tz)^2(ty)$$
$$= t^3(xy^2 + x^2y + y^3 - y^2y) = t^3R(x,y,z) \quad \triangleright$$

Una relación $R(x,y,z)$ homogénea, define una superficie como se muestra en el siguiente Lema.

LEMA 6.1 *La relación $R(x,y,z) = 0$, donde $R(x,y,z)$ es homogénea de grado s define una superficie cónica.*

$\triangleleft$ Sea $R(x,y,z) = 0$ tal que R es homogénea de grado s, y supóngase que $v = (0,0,0)$. Sea $p_0 = (x_0,y_0,z_0)$ tal que $R(y_0,x_0,z_0) = 0$.

Se afirma que toda la recta por p_0 y v satisface la relación. Tal recta tiene, en las variables x,y, la ecuación

$$\begin{cases} x = tx_0 \\ y = ty_0 \\ z = tz_0 \end{cases}$$

De esta manera, si el punto (x,y,z) está sobre la recta, entonces

$$R(x,y,z) = R(tx_0,ty_0,tz_0) = t^sR(x_0,y_0,z_0) = t^s(0) = 0$$

lo que demuestra que $(x,y,z) \in S$. $\triangleright$

EJEMPLO. La relación

$$R(x,y,z) = x^2 + y^2 - z^2$$

es homogénea de grado 2 y por lo tanto, la relación

$$x^2 + y^2 - z^2 = 0$$

define una superficie cónica.

$\triangleleft$ Claramente $R(0,0,0) = 0^2 + 0^2 - 0^2 = 0$, lo que nos dice que el punto $(0,0,0)$ está en la superficie y es su vértice.

De la relación $x^2 + y^2 - z^2 = 0$, si despejamos a z^2, obtenemos

$$x^2 + y^2 = z^2.$$



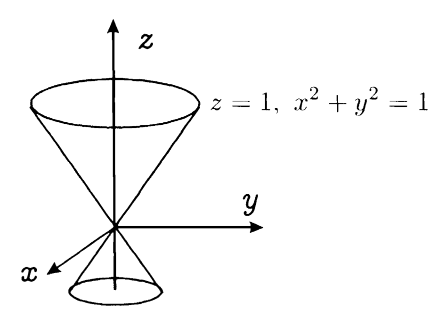


Figura 6.10: Cono circular.

Si tomamos $z = 1$, entonces $x^2 + y^2 = 1$ es un círculo de radio 1, lo que nos dice que la curva γ, base de la superficie es un círculo.

La superficie cónica S es un **cono circular** con vértice en $(0, 0, 0)$. La figura 6.10 muestra esta superficie. $\triangleright$

EJEMPLO. Sea la relación

$$R(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2}$$

homogénea de grado 2, entonces por el Lema 6.1, define una superficie cónica.

$\triangleleft$ Al hacer $R(x, y, z) = 0$ se obtiene la ecuación

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$$

o bien, despejando los términos en z se tiene la relación equivalente

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}$$

Si se hace $z = c$ obtenemos una curva γ base de la superficie, dada por

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

que es la ecuación de una elipse con parámetros a, b, a una altura $z = c$.

Le llamaremos a tal superficie un **cono elíptico** con vértice en el punto $v = (0, 0, 0)$, y se ilustra en la figura 6.11. $\triangleright$



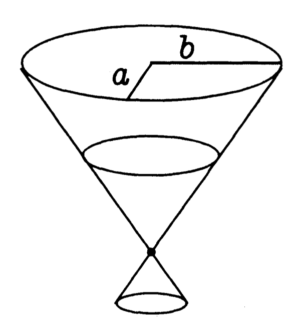


Figura 6.11: Cono elíptico.

6.4 Elipsoides, Hiperboloides y Paraboloides

En esta parte definimos nuevas superficies a partir de las ya conocidas, utilizando **cambios afines** de coordenadas, como la deformación de los ejes cordenados y las traslaciones.

◁ Dada la esfera unitaria en $\mathbb{R}^3$,

$$x^2 + y^2 + z^2 = 1,$$

en las coordenadas cartesianas x, y, z, la podemos deformar en un **elipsoide** en $\mathbb{R}^3$,

$$\frac{u^2}{a^2} + \frac{v^2}{b^2} + \frac{w^2}{c^2} = 1$$

en las nuevas coordenadas u, v, w, mediante la deformación de los ejes coordenados

$$\begin{cases} x = \frac{u}{a} \\ y = \frac{v}{b} \\ z = \frac{w}{c} \end{cases},$$

donde los parámetros $a, b, c > 0$, o bien, mediante las relaciones inversas,

$$\begin{cases} u = ax \\ v = by \\ w = cz \end{cases}$$

lo cual se comprueba con la simple sustitución

$$1 = x^2 + y^2 + z^2 = \left(\frac{u}{a}\right)^2 + \left(\frac{v}{b}\right)^2 + \left(\frac{w}{c}\right)^2 = \frac{u^2}{a^2} + \frac{v^2}{b^2} + \frac{w^2}{c^2}.$$

La figura 6.12 ilustra al elipsoide. ▷

Mediante procesos análogos de deformación de los ejes y usando otras cordenadas[1], es posible obtener superficies nuevas a partir de superficies ya conocidas. A continuación damos ejemplos.

EJEMPLO. (Elipsoide circular). Sea γ la elipse centrada en $(0,0)$ del plano $\mathbb{R}^2_{x,z}$, definida por la ecuación

$$\frac{x^2}{a^2} + \frac{z^2}{c^2} = 1, \quad a, c > 0$$

$\triangleleft$ Si se hace girar γ alrededor del eje z, al reemplazar $x \to \sqrt{x^2 + y^2}$ se obtiene la superficie de revolución elipsoidal dada por la ecuación

$$1 = \frac{(\sqrt{x^2 + y^2})^2}{a^2} + \frac{z^2}{c^2} = \frac{x^2 + y^2}{a^2} + \frac{z^2}{c^2} = \frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{c^2}$$

llamada **elipsoide circular** (véase la figura 6.12).

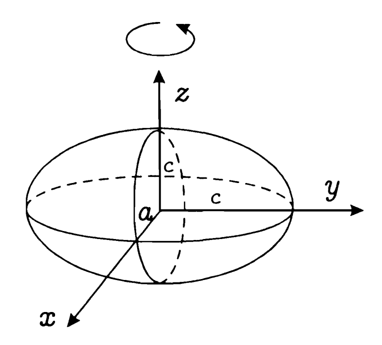


Figura 6.12: Elipsoide circular.

Al hacer una deformación del eje y mediante el reemplazo de la variable y por

$$y \to \frac{ay}{b}$$

y al sustituir en la última ecuación se tiene

$$1 = \frac{x^2}{a^2} + \frac{\left(\frac{ay}{b}\right)^2}{a^2} + \frac{z^2}{c^2} = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}$$

es decir,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$



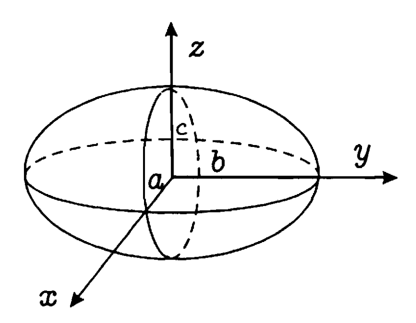


Figura 6.13: Elipsoide.

que es la ecuación de un **elipsoide** con parámetros a, b, c (véase la figura 6.13). ▷

EJEMPLO. (Hiperboloide de una hoja). Consideremos la hipérbola en el plano $\mathbb{R}^2_{x,z}$ dada por la ecuación,

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1$$

donde los parámetros $a, b > 0$.

◁ Al hacerla girar tal curva alrededor del eje z, reemplazando $x \to \sqrt{x^2 + y^2}$ se obtiene una superficie denominada **hiperboloide circular de una hoja** dado por la relación,

$$1 = \frac{(\sqrt{x^2 + y^2})^2}{a^2} - \frac{z^2}{c^2} = \frac{x^2 + y^2}{a^2} - \frac{z^2}{c^2}$$

es decir,

$$1 = \frac{x^2}{a^2} + \frac{y^2}{a^2} - \frac{z^2}{c^2}$$

Como cada punto de la hipérbola genera un círculo al girarla. Esto se ilustra en la figura 6.14.

Si se deforma el eje y con el reemplazo

$$y \to \frac{ay}{b}$$

se obtiene la nueva relación

$$1 = \frac{x^2}{a^2} + \frac{\left(\frac{ay}{b}\right)}{a^2} - \frac{z^2}{c^2} = \frac{x^2}{a^2} + \frac{\frac{a^2y^2}{b^2}}{a^2} - \frac{z^2}{c^2} = \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2}$$

[1]En la sección 5.4 damos un tratamiento más completo a éste proceso de cambiar coordenadas en el espacio $\mathbb{R}^3$.



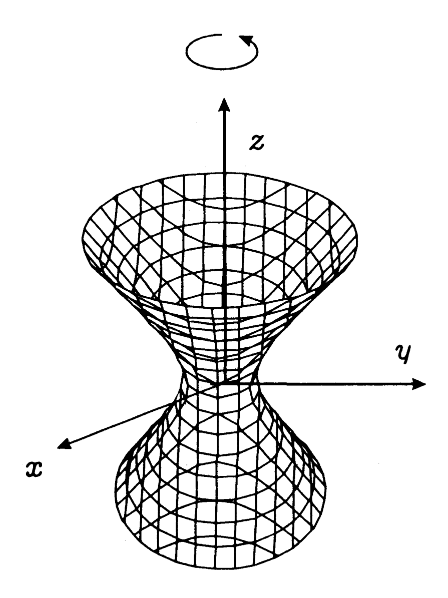


Figura 6.11: Hiperboloide circular de una hoja.

o equivalentemente,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

que corresponde a una **hiperboloide elíptico de una hoja** (véase la figura 6.15).

Para muestra, la relación

$$2x^2 + \frac{1}{2}y^2 - 3z^2 = 1$$

define un hiperboloide elíptico de una hoja ▷

EJEMPLO. (Hiperboloide de dos hojas). Si la misma hipérbola del ejemplo precedente en el plano $\mathbb{R}^2_{x,z}$ con la ecuación

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1$$

se hace girar alrededor del eje x, al reemplazar $z \to \sqrt{y^2 + z^2}$ se obtiene la superficie de revolución con ecuación

$$1 = \frac{x^2}{a^2} - \frac{(\sqrt{y^2 + z^2})^2}{c^2} = \frac{x^2}{a^2} - \frac{y^2 + z^2}{c^2}$$



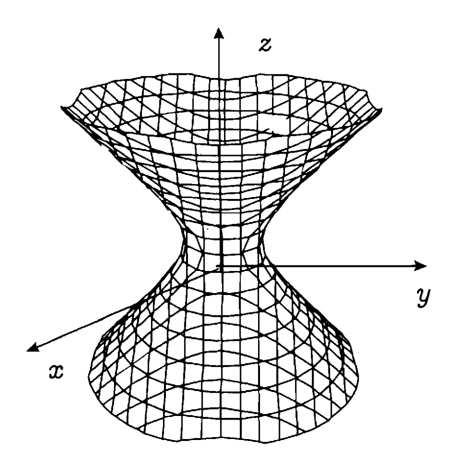


Figura 6.15: Hiperboloide elíptico de una hoja.

es decir,

$$1 = \frac{x^2}{a^2} - \frac{y^2}{c^2} - \frac{z^2}{c^2}$$

y tal superficie se llamará **hiperboloide circular de dos hojas**.

Al hacer la deformación del eje y mediante el cambio coordenado

$$y \to \frac{cy}{b}$$

se obtiene la nueva ecuación

$$1 = \frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2}$$

que corresponde a un **hiperboloide elíptico** de dos hojas (véase la figura 6.16).

Para muestra, la ecuación

$$-x^2 - y^2 + \frac{1}{2}z^2 = 1$$

es un **hiperboloide elíptico** de dos hojas. ▷

EJEMPLO. (Paraboloide elíptico). Sea la parábola cuadrática en el plano $\mathbb{R}^2_{x,z}$ dada por la ecuación

$$az = x^2$$

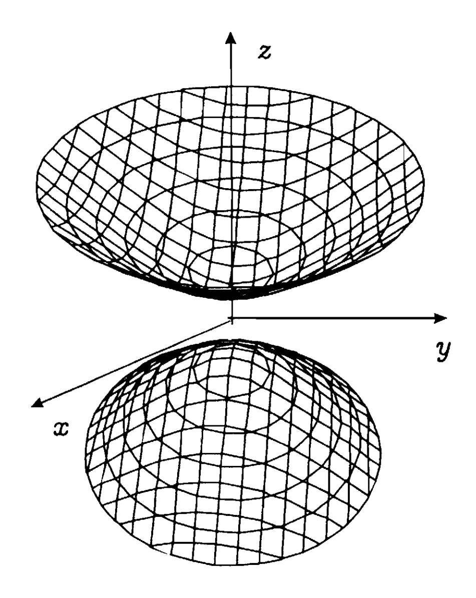


Figura 6.16: Hiperboloide elíptico de dos hojas.

◁ Si se hace girar alrededor del eje z haciendo el reemplazo canónico $x \to \sqrt{x^2 + y^2}$ en la última relación, se obtiene un **paraboloide** con ecuación

$$az = \left(\sqrt{x^2 + y^2}\right)^2 = x^2 + y^2$$

o equivalentemente,

$$z = \frac{x^2 + y^2}{a} = \frac{x^2}{a} + \frac{y^2}{a}$$

o bien,

$$z = \frac{x^2}{a} + \frac{y^2}{a}$$

llamado **paraboloide circular**.

Si se deforma el eje y por el cambio coordenado $y \to \frac{\sqrt{a}y}{\sqrt{b}}$ se obtiene nueva la ecuación

$$z = \frac{x^2}{a^2} + \frac{\left(\frac{\sqrt{a}y}{\sqrt{b}}\right)^2}{a} = \frac{x^2}{a} + \frac{\frac{ay^2}{b}}{a} = \frac{x^2}{a} + \frac{y^2}{b}$$

es decir,

$$z = \frac{x^2}{a} + \frac{y^2}{b}$$

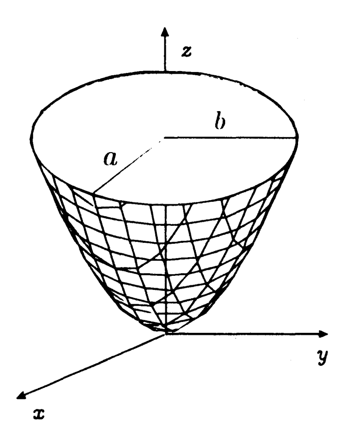


Figura 6.17: Paraboloide elíptico.

con los parámetros a, b, con el mismo signo, que representa un **paraboloide elíptico** (véase la figura 6.17). ▷

NOTA. Si en la ecuación obtenida en el ejemplo anterior

$$z = \frac{x^2}{a} + \frac{y^2}{b}$$

se tienen signos diferentes para a y b, la superficie se llamará **paraboloide hiperbólico**, y su gráfica se muestra en la figura 6.18. A esta superficie se le conoce también como **silla de montar**.

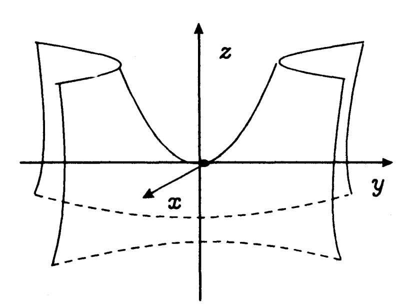


Figura 6.18: Paraboloide hiperbólico.

Observamos que hasta este punto todas nuestras superficies obtenidas mediante la deformación de los ejes han estado centradas en el origen de

coordenadas y todas tienen una descripción dada por una fórmula cuadrática. En realidad, si se tiene una expresión cuadrática del tipo

$$Ax^2 + By^2 + Cz^2 + Dx + Ey + Fz + G = 0$$

donde no hay términos cruzados en las variables, [2] siempre es posible identificarla como una de las superficies de la lista que hemos mencionado. El proceso para la identificación involucra una **translación de ejes coordenados** del origen en $\mathbb{R}^3$ a otro punto del espacio.

Si $p = (h, k, l)$ es un punto del espacio $\mathbb{R}^3$, entonces al utilizar las coordenadas (u, v, w) en $\mathbb{R}^3$ ligadas a las coordendas cartesianas mediante las fórmulas

$$\left\{\begin{array}{c} x = u + h \\ y = v + k \\ z = w + l \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{c} u = x - h \\ v = y - k \\ w = z - l \end{array}\right\}$$

conseguimos una translación del punto origen $(0, 0, 0)$ al punto $p = (h, k, l)$ como lo muestra un cálculo simple (de hecho una sustitución).

Si $x = h$, $y = k$, $z = l$, entonces $u = 0$, $v = 0$, $w = 0$, y reciprocamente si $u = 0$, $v = 0$, $w = 0$, entonces $x = h$, $y = k$, $z = l$, y como lo muestra la figura 6.19.

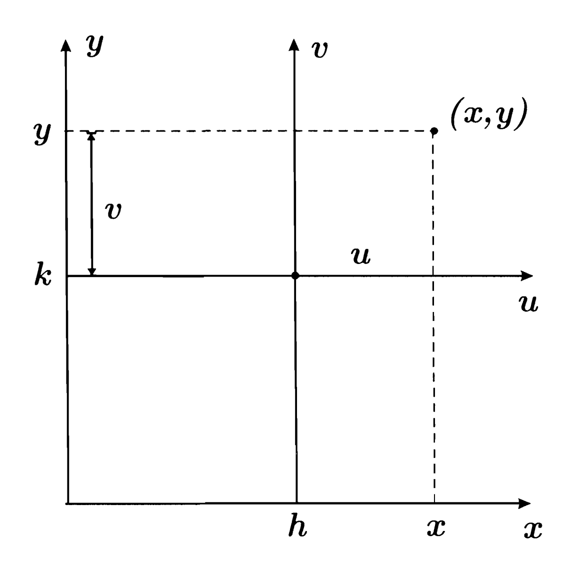


Figura 6.19: Translación de ejes coordenados.

[2]Si hay términos cruzados el proceso de identificación se complica y para ello es necesario utilizar métodos del **Algebra Lineal**.

EJEMPLO. Hacer una translación de ejes del origen de coordenadas al punto $p = (-1, 3, -2)$.

◁ Al utilizar fórmulas se tiene que las nuevas variables u, v, w en este sistema nuevo son dadas por,

$$\left\{ \begin{array}{c} u = x - (-1) = x + 1 \\ v = y - 3 = y - 3 \\ w = z - (-2) = z + 2 \end{array} \right\} \Longleftrightarrow \left\{ \begin{array}{c} x = u - 1 \\ y = v + 3 \\ z = w - 2 \end{array} \right\} \quad \triangleright$$

EJEMPLO. Identificar a la superficie espacial dada por la ecuación

$$x^2 + y^2 + z^2 - x + y = 0$$

◁ Hacemos un completamiento de términos para tener trinomios cuadrados perfectos, asociando los semejantes,

$$0 = x^2 + y^2 + z^2 - x + y = (x^2 - x) + (y^2 + y) + z^2$$

de donde,

$$\left(x^2 - 2\left(\frac{1}{2}\right)x + \left(\frac{1}{2}\right)^2\right) + \left(y^2 + 2\left(\frac{1}{2}\right)y + \left(\frac{1}{2}\right)^2\right) + z^2 = \frac{1}{4} + \frac{1}{4}$$

o

$$\left(x - \frac{1}{2}\right)^2 + \left(y + \frac{1}{2}\right)^2 + (z - 0)^2 = \frac{1}{2}$$

Al hacer la translación de ejes con nuevas coordenadas u, v, w dadas por

$$\left\{ \begin{array}{c} u = x - 1/2 = x - 1/2 \\ v = y + 1/2 = y - (-1/2) \\ z = z - 0 = z - 0 \end{array} \right.$$

tenemos el nuevo origen en el punto $(1/2, -1/2, 0)$ y la ecuación en estas nuevas coordenadas dada por

$$u^2 + v^2 + z^2 = \frac{1}{2}$$

que es una esfera de radio $\sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2}$

De esta manera, la ecuación inicial describe a una esfera de radio $R = \frac{\sqrt{2}}{2}$ con centro en el punto $c = \left(\frac{1}{2}, \frac{-1}{2}, 0\right)$. ▷

EJEMPLO. Identificar a la superficie en $\mathbb{R}^3$ dada por la ecuación

$$4x^2 + 4z^2 - y^2 - 8x + 8z + 4y + 4 = 0$$



◁ Al agrupar términos obtenemos que

$$\begin{aligned} 0 &= (4x^2 - 8x) - (y^2 - 4y) + (4z^2 + 8z) + 4 \\ &= 4(x^2 - 2x) - (y^2 - 4y) + 4(z^2 + 2z) + 4 \\ &= 4(x^2 - 2x + 1) - (y^2 - 4y + 4) + 4(z^2 + 2z + 1) + 4 - 4 \end{aligned}$$

o bien,

$$4(x-1)^2 - (y+2)^2 + 4(z+1)^2 = 0$$

Al hacer la translación de ejes al punto $(1, -2, -1)$ mediante las ecuaciones

$$\begin{cases} u = x - 1 \\ v = y + 2 \\ w = z + 1 \end{cases} \iff \begin{cases} x = u + 1 \\ y = v - 2 \\ z = w - 1 \end{cases}$$

obtenemos la nueva ecuación en el sistema de coordenadas u, v, w dado por

$$4u^2 - v^2 + 4w^2 = 0$$

Ya que la relación es homogénea de grado 2, se tiene una superficie cónica, y al despejar la variable con signo negativo, se tiene

$$v^2 = 4u^2 + 4w^2$$

De esta forma, si $v = c$ es una constante arbitraria se obtiene la ecuación de la base del cono

$$c^2 = 4u^2 + 4w^2 \iff \left(\frac{c}{2}\right)^2 = u^2 + w^2$$

que es un círculo de radio $R = c/2$.

De esta manera, la ecuación

$$\triangleleft \quad 4x^2 + 4z^2 - y^2 + 8x + 8z + 4y + 4 = 0$$

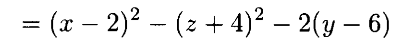
define a un cono circular en dirección del eje y con vértice en el punto $(1, -2, -1)$. ▷

EJEMPLO. Identificar la superficie dada por la ecuación

$$x^2 - z^2 - 4x + 8z - 2y = 0$$

◁ Al completar cuadrados obtenemos

$$0 = (x^2 - 4x) - (z^2 - 8z) - 2y = (x^2 - 4x + 4) - (z^2 - 8z + 16) - 2y + 12$$



$$= (x-2)^2 - (z+4)^2 - 2(y-6)$$

de donde,

$$2(y-6) = (x-2)^2 - (z+4)^2$$

o bien,

$$(y-6) = \frac{(x-2)^2}{2} + \frac{(z+4)^2}{-2}$$

Al realizar la translación de ejes al punto $(2, 6, -4)$ usando las ecuaciones

$$\begin{cases} u = x - 2 \\ v = y - 6 \\ w = z + 4 \end{cases} \Longleftrightarrow \begin{cases} x = u + 2 \\ y = v + 6 \\ z = w - 4 \end{cases}$$

se obtiene la ecuación

$$v = \frac{u^2}{2} + \frac{w^2}{-2}$$

que corresponde a un paraboloide hiperbólico.

De esta manera, la ecuación inicial describe a un paraboloide hiperbólico centrado en el punto $(2, 6, -4)$. $\triangleright$

EJEMPLO. Analizar la ecuación cuadrática

$$9x^2 + 4y^2 - 36z^2 - 18x - 8y - 72z + 13 = 0$$

e identificar a la superficie que define.

$\triangleleft$ De la completación de cuadrados en tal ecuación se obtiene

$$0 = (9x^2 - 18x) + (4y^2 + 8y) - (36z^2 + 72z) + 13$$

$$= 9(x^2 - 2x) + 4(y^2 + 2y) - 36(z^2 + 2z) + 13$$

$$= 9(x^2 - 2x + 1) + 4(y^2 + 2y + 1) - 36(z^2 + 2z + 1) + 13 - 9 - 4 + 36$$

es decir,

$$-36 = 9(x-1)^2 + 4(y+1)^2 - 36(z+1)^2$$

o bien,

$$1 = \frac{-9}{36}(x-1)^2 - \frac{4}{36}(y+1)^2 + (z+1)^2$$

lo que equivale a

$$1 = (z+1)^2 - \frac{(x-1)^2}{4} - \frac{(y+1)^2}{9}$$



al efectuar la translación de ejes al punto $(1, -1-1)$ mediante al cambio de coordenadas

$$\left\{\begin{array}{l} u = x - 1 \\ v = y + 1 \\ w = z + 1 \end{array}\right. \iff \left\{\begin{array}{l} x = u + 1 \\ y = v - 1 \\ z = w - 1 \end{array}\right.$$

obtenemos una nueva ecuación

$$1 = w^2 - \frac{u^2}{4} - \frac{v^2}{9}$$

que corresponde a un hiperboloide elíptico de dos hojas.

De esta manera, la ecuación inicial define a un hiperboloide elíptico de dos hojas concentrado en el punto $(1, -1, -1)$. ▷

Ejercicios.

1. Dar las coordendas del centro y dar el radio de la esfera definida por la ecuación

a. $x^2 + y^2 + z^2 - y + 2z + 1 = 0$
b. $x^2 + y^2 + z^2 = 4z - 3$
c. $2x^2 + 2y^2 + 2z^2 = 3z - 4y - 2$

2. Dado el punto $p = (0, -1, 2)$ decir si está dentro, fuera o sobre cada una de las esferas dadas en los incisos del ejercicio 1.

3. Decir qué superficies espaciales definen las siguientes ecuaciones

a. $y^2 = 2x$
b. $z^2 = xz$
c. $y^3 = 4x^2$
d. $x^2 + y^2 = 4$
e. $x^2 - y^2 = 1$
f. $x^2 - z^2 = 0$

4. Dar la ecuación del cono cuyo vértice es el punto $(0, 0, -1)$ y cuya base es la curva

a. elíptica $\frac{x^2}{4} + \frac{y^2}{9} = 1$
b. circular $4x^2 + 4y^2 = 1$

5. Calcular la ecuación de la superficie S obtenida al hacer girar

a. La recta $x + y = 1$ en el plano $\mathbb{R}^2_{x,y}$ alrededor del eje x.
b. La parábola $y = z^2$ en el plano $\mathbb{R}^2_{y,z}$ alrededor del eje y.
c. La hipérbola $xz = 1$ en el plano $\mathbb{R}^2_{x,z}$ alrededor del eje z.



6. a. Dar las ecuaciones de las líneas de intersección de la superficie dada por $z = y^2 - x^2$ con los planos $z = 1$, $x = 1$, $y = 1$, respectivamente.
b. Dar la ecuación de la línea de intersección de las superficies

$$\begin{cases} z = 1 - x^2 - y^2 \\ z = x^2 + y^2 \end{cases}$$

7. Identificar a las siguientes superficies mediante una translación de ejes.
a. $x^2 + z^2 - 4x - 4z + 4 = 0$
b. $x^2 + y^2 - z^2 - 2y + 2z = 0$
c. $x^2 + 2y^2 + 2z^2 - 4y + 4z + 4 = 0$
d. $4x^2 + y^2 - z^2 - 24x - 4y + 2z + 35 = 0$
e. $x^2 + y^2 - z^2 - 2x - 2y + 2z + 2 = 0$
f. $x^2 + y^2 - 6x + 6y - 4z + 18 = 0$
g. $9x - z^2 - 18x - 18y - 6z = 0$


Casa abierta al tiempo

Capítulo 7

Orientación de curvas y poligonales

7.1 Orientación.

En este capítulo introducimos los conceptos de **orientación** de una curva suave y el de curva **poligonal** (curva suave por pedazos), que resulta de gran utilidad para las aplicaciones en las ciencias naturales.

Dada una curva $\Gamma \subset \mathbb{R}^3$ parametrizada por la función

$$\gamma : [a, b] \to \mathbb{R}^3$$

de tal forma que $||\dot{\gamma}(t)|| \neq 0$, diremos que la **orientación** de Γ es del punto $\gamma(a) \in \Gamma$ al punto $\gamma(b) \in \Gamma$, y que está inducida por la orientación del intervalo orientado $[a, b]^+$ (véase la figura 7.1).

Distinguimos a la curva orientada de esta forma por Γ^+. Naturalmente, esto hace que Γ^+ esté orientada por el vector tangente $\dot{\gamma}(t)$.

Supongamos que el parámetro t es una función de un nuevo parámetro $\tau \in [c, d]$, esto es $t = t(\tau)$, y que es una biyección diferenciable

$$t : [c, d] \to [a, b]$$

Entonces diremos que Γ se puede reparametrizar por el parámetro τ en el intervalo $[c, d]$ por la composición

$$\widetilde{\gamma}(\tau) = \gamma(t(\tau))$$

Dicho de otra forma, Γ está parametrizada por la función $\widetilde{\gamma} : [c, d] \to \mathbb{R}^3$. La función $\widetilde{\gamma}$ se dice ser una **reparametrización** de Γ.

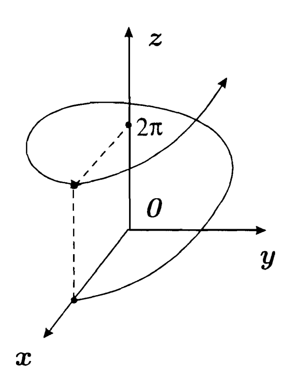


Figura 7.1: Hélice unitaria de paso 2π

EJEMPLO. Sea Γ la primera espira de la hélice unitaria de paso 2π, con la parametrización

$$\gamma(t) = (\cos t,\, \operatorname{sen} t, t)$$

en el intervalo $[0, 2\pi]$. Esto orienta a Γ^+ del punto $(1, 0, 0)$ al punto $(1, 0, 2\pi)$.

$\lhd$ Sea ahora la reparametrización de Γ inducida por el cambio de parámetro

$$t = 2\pi - \tau$$

definida para $\tau \in [0, 2\pi]$,

$$\widetilde{\gamma}(\tau) = (\cos t(2\pi - \tau),\, \operatorname{sen}(2\pi - \tau), 2\pi - \tau)$$

En este caso $\widetilde{\gamma}$ invierte la orientación de Γ pues $\widetilde{\gamma}(0) = (1, 0, 2\pi)$ y $\widetilde{\gamma}(2\pi) = (1, 0, 0)$.

Notamos que la función biyectiva que cambia los parámetros

$$t : [0, 2\pi] \to [0, 2\pi], \quad t = 2\pi - \tau$$

es tal que $\frac{dt}{d\tau} = -1 < 0$. La figura 7.1 ilustra esta situación. $\rhd$

La última observación en el ejemplo anterior no es extraña. De hecho, el siguiente resultado nos impone condiciones sobre la derivada de una función que cambia los parámetros, para que en la reparametrización se conserve la orientación.

LEMA 7.1 *Sea $\gamma : [a, b] \to \mathbb{R}^3$ una parametrización de la curva Γ que induce una orientación Γ^+. Si $t \in [a, b]$ es el parámetro de γ y existe un*

cambio de parámetro $t = t(\tau)$ definido en el intervalo $[c, d]$, entonces la reparametrización dada por

$$\widetilde{\gamma}(\tau) = \gamma(t(\tau))$$

es tal que,

a. *conserva la orientación de Γ, si $\frac{dt}{d\tau} > 0$ en todo el intervalo $[c, d]$.*

b. *cambia la orientación de Γ, si $\frac{dt}{d\tau} < 0$ en todo el intervalo $[c, d]$.*

$\lhd$ Por la regla de la cadena, se tiene que

$$\frac{d}{d\tau}\widetilde{\gamma}(\tau) = \dot{\gamma}(t)\frac{dt}{d\tau}$$

lo que nos indica que los vectores $\frac{dt}{d\tau}\widetilde{\gamma}$ y $\dot{\gamma}$ son paralelos. La cantidad $\frac{dt}{d\tau}$ indica el sentido de cada uno y, por lo tanto, la orientación . $\rhd$

A una curva Γ que es la unión (+) de las curvas suaves $\Gamma_1, \Gamma_2, \cdots, \Gamma_k$ se le llamará una **curva poligonal orientada**, o **curva suave por pedazos** orientada (véase la figura 7.2 a.)

La definimos por simplicidad por

$$\Gamma^+ = \Gamma_1^+ + \Gamma_2^+ + \cdots + \Gamma_k^+$$

imponiendo la condición de que la curva Γ_j^+ inicia donde termina Γ_{j-1}^+.

EJEMPLO. Sea Γ la unión (+) de los segmentos de las parábolas $y_1 = x^2$, $y_2 = 2 - x^2$ comprendidos entre sus intersecciones en los puntos $(1,1)$ y $(-1,1)$.

$\lhd$ Si Γ_1 es el segmento de la parábola $y = x^2$ en el intervalo $[-1,1]$ y Γ_2 es el segmento de la parábola $y = 2 - x^2$ en el intervalo $[-1,1]$, entonces $\Gamma^+ = \Gamma_1^+ + \Gamma_2^+$ es tal que la orientación inicia en el punto $(-1,1)$, recorre por Γ_1^+ hasta $(1,1)$ y luego recorre por Γ_2^+ hasta $(-1,1)$ siguiendo la orientación invertida del intervalo $[-1,1]$ que definimos por $[-1,1]^-$. La curva Γ^+ es en total una curva cerrada iniciándose en el punto $(-1,1)$ y cerrando su circuito en el mismo punto (véase la figura 7.2 b.)

Damos una parametrización de $\Gamma^+ = \Gamma_1^+ + \Gamma_2^2$ como sigue.

Γ_1 es la curva definida por $y = x^2$ si $x \in [-1,1]^+$, de donde, poniendo $x = t$ y $y = t^2$ se obtiene la curva $\gamma_1(t) = (t, t^2)$ con $t \in [-1,1]^+$ que da orientación a Γ_1^+.

Para parametrizar a Γ_2 con la orientación que define a la poligonal Γ^+, siendo el segmento inmediato de Γ_1^+, consideramos su definición. Γ_2 es la curva $y = 2 - x^2$ con $x \in [-1,1]$, lo que indica que Γ_2^+ está orientada por el intervalo $[-1,1]^-$. Usemos el cambio lineal de intervalos

$$x : [1,2]^+ \rightarrow [-1,1]^- = [1,-1]^+$$

que lleva $1 \to 1$, $2 \to -1$ definido por,

$$x = x(t) = -2(t-1) + 1 = 3 - 2t.$$

Ya que los puntos sobre la curva Γ_2 son de la forma $(x, 2-x^2)$, entonces $\gamma_2(t) = (3-2t,\ 2-(3-2t)^2)$ con $t \in [1,2]$ parametriza a Γ_2^+. Por lo tanto, la curva $\gamma = \gamma(t)$ que parametriza a todo el circuito con la orientación seguida de $\Gamma^+ = \Gamma_1^+ + \Gamma_2^+$ es

$$\gamma(t) = \begin{cases} (t, t^2) & -1 \leq t \leq 1 \\ (3-2t, 2-(3-2t)^2) & 1 \leq t \leq 2 \end{cases}$$

donde $\gamma : [-1,2] \to \mathbb{R}^2$ es regular por pedazos ▷

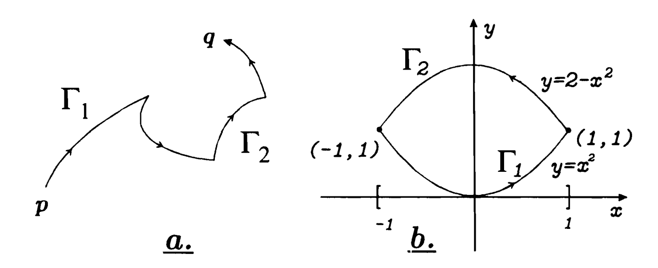


Figura 7.2: a. Poligonal b. Curva Γ cerrada orientada.

El ejemplo anterior nos da una manera de orientar a una región plana Ω que tiene como frontera a una curva cerrada poligonal Γ. Para orientar a tal región, parametrizamos a Γ de tal forma que recorra el circuito en sentido contrario a las manecillas del reloj (**regla de la mano derecha**)

Esto es equivalente a considerar el **marco** dado por la pareja ordenada de vectores $\{n(t), \dot{\gamma}(t)\}$, donde $n(t)$ es un **vector normal** a $\gamma(t)$ y $\dot{\gamma}(t)$ es el vector tangente, satisfaciendo la relación de orientación

$$0 < det \begin{pmatrix} n(t) \\ \dot{\gamma}(t) \end{pmatrix} = -det \begin{pmatrix} \dot{\gamma}(t) \\ n(t) \end{pmatrix}$$

No es difícil ver que esto se logra al orientar la curva cerrada Γ^+ en contra de las manecillas del reloj y tomar sus vectores normales $n(t)$ apuntando hacia afuera de Ω (véase la figura 7.3). Esta es una manera de orientar a una región Ω provista con una frontera conformada por una poligonal cerrada. La definiremos por Ω^+.



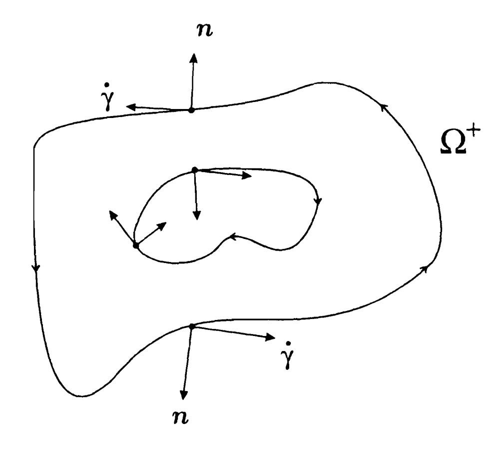


Figura 7.3: Orientación de una región plana Ω^+.

7.2 Longitud de arco y ángulo entre curvas

Ahora se exponen los conceptos de **longitud de arco** de curva suave y el de **ángulo** entre dos curvas regulares.

Dada una curva regular $\Gamma \subset \mathbb{R}^3$, parametrizada por

$$\gamma : [a, b] \to \Gamma,$$

si se toma $p \in \Gamma$, hasta un infinitésimo de longitud $||d\gamma||$ en tal punto de la curva, éste se descompone infinitesimalmente como

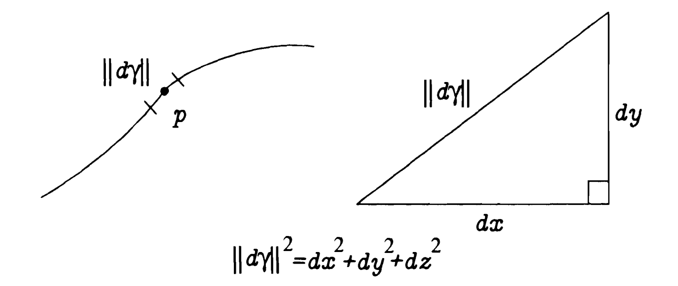


Figura 7.4: Longitud de arco de una curva.

$$||d\gamma||^2 = dx^2 + dy^2 + dz^2,$$

donde dx, dy, dz son las componentes infinitesimales del infinitésimo del arco (vectorial) $d\gamma$ (véase la figura 7.4).

De allí que
$$||d\gamma|| = \sqrt{dx^2 + dy^2 + dz^2}$$

De esta forma, si $\ell(\Gamma)$ define la longitud de Γ entonces, esta se calcula por
$$\ell(\Gamma) = \int_\Gamma ||d\gamma|| = \int_\Gamma \sqrt{dx^2 + dy^2 + dz^2}$$

Por lo tanto, si en cordenadas cartesianas $\gamma = \gamma(t)$ parametriza a Γ en el intervalo $[a, b]$,
$$\gamma(t) = (x(t),\, y(t),\, z(t)),$$

entonces
$$\begin{cases} dx = \dot{x}(t)\, dt & \\ dy = \dot{y}(t)\, dt & t \in [a, b] \\ dz = \dot{z}(t)\, dt, & \end{cases}$$

y de esta manera, la longitud de Γ se calcula por

$$\ell(\Gamma) = \int_\Gamma \sqrt{dx^2 + dy^2 + dz^2} = \int_{[a,b]} \sqrt{\dot{x}(t)^2 dt^2 + \dot{y}(t)^2 dt^2 + \dot{z}(t)^2 dt^2}$$

$$= \int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2} dt.$$

La igualdad
$$\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2 = ||v(t)||^2,$$

hace que la siguiente definición sea natural.

DEFINICIÓN. La **longitud** de la curva regular parametrizada $\Gamma \subset \mathbb{R}^3$ se define por

$$\ell(\Gamma) = \int_a^b \sqrt{< v(t),\, v(t) >}\, dt = \int_a^b ||v(t)|| dt,$$

donde $v(t) = \dot{\gamma}(t)$ es el vector velocidad tangente a la curva Γ en el punto $\gamma(t)$.

El siguiente ejemplo ilustra esta definición que se ha heredado directamente de los resultados básicos del cálculo.

EJEMPLO. Sea el subconjunto del plano

$$S_R^1 = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = R\}$$

identificado como el **círculo** de radio R y centro en $(0, 0)$.



◁ Una parametrización que hace de S^1_R una curva regular, está dada por el sistema

$$\gamma(t):\begin{cases} x = R\cos t \\ y = R\,\text{sen}t \end{cases}, \quad t \in [0, 2\pi]$$

De la definición, la longitud de S^1_R se calcula por

$$\ell = \ell(S^1_R) = \int_0^{2\pi} \sqrt{R^2\,\text{sen}^2 t + R^2\cos^2 t}\,dt$$

$$= \int_0^{2\pi} \sqrt{R^2(\text{sen}^2 t + \cos^2 t)}dt = \int_0^{2\pi} Rdt = 2\pi R. \quad \triangleright$$

EJEMPLO. Sea $\Gamma \subset \mathbb{R}^3$ la hélice contenida en el cilindro de radio 4 y paso de altura $h = 4\pi$.

◁ Ya se vió en el Capítulo 3, que las ecuaciones que definen paramétricamente a Γ son

$$\begin{cases} x = 4\cos t \\ y = 4\,\text{sen}\,t \\ z = 2t \end{cases}$$

lo que implica que,

$$||v(t)||^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2 = 16\text{sen}^2 t + 16\cos^2 t + 4 = 20,$$

debido a que $v(t) = (-4\,\text{sen}\,t,\, 4\cos t, 2)$.

Por lo tanto, en el intervalo $[0, 2\pi]$, la longitud de esta hélice, hasta el primer paso es

$$\ell = \int_0^{2\pi} ||v(t)||dt = \int_0^{2\pi} \sqrt{20}dt = 2\sqrt{20}\pi \quad \triangleright$$

El siguiente resultado nos asegura la independencia de la longitud de una curva respecto a una parametrización. Esto es, la definición de longitud está bien definida, pues no depende de la parametrización escogida de la curva.

LEMA 7.2 *Sea Γ una curva regular parametrizada por*

$$\gamma : [a, b] \to R$$

donde $\gamma = \gamma(t)$, $t \in [a, b]$. Si $\tilde{\gamma} : [c, d] \to \Gamma$, con $\bar{\gamma} = \bar{\gamma}(\tau)$ es otra parametrización tal que el parámetro temporal $t = t(\tau)$ es una función del parámetro $\tau \in [c, d]$,

$$t : [c, d] \to [a, b],$$

con $\frac{dt}{d\tau} > 0$, entonces las longitudes del arco en las parametrizaciones respectivas son las mismas, esto es,

$$\ell = \int_a^b ||\dot{\gamma}(t)||dt = \int_c^d ||\dot{\tilde{\gamma}}(\tau)||d\tau = \tilde{\ell}$$

◁ Defínase

$$\tilde{\gamma}(\tau) = (x^1(t(\tau)), x^2(t(\tau)), x^3(t(\tau)) = (y^1(\tau), y^2(\tau), y^3(\tau)),$$

y sea el vector velocidad

$$\dot{\tilde{\gamma}} = w(\tau) = \left(\frac{dy^1}{d\tau}, \frac{dy^2}{d\tau}, \frac{dy^3}{d\tau}\right), \quad c \leq \tau \leq d.$$

Entonces, por definición,

$$\tilde{\ell} = \int_c^d ||w(\tau)||d\tau.$$

Por otro lado, por La regla de la cadena,

$$||w(\tau)|| = \sqrt{\sum_{i=1}^{3}\left(\frac{dy^i}{d\tau}\right)^2} = \sqrt{\sum_{i=1}^{3}\left(\frac{dx^i}{dt}\frac{dt}{d\tau}\right)^2}$$

$$= \left|\frac{dt}{d\tau}\right| \sqrt{\sum_{i=1}^{3}\left(\frac{dx^i}{dt}\right)^2} = \left|\frac{dt}{d\tau}\right| \, ||v(t)||,$$

De esta forma, como $dt/d\tau > 0$, por el teorema de cambio de variable se tiene que

$$\tilde{\ell} = \int_c^d ||w(\tau)||d\tau = \int_c^d ||v(t(\tau))||\frac{dt}{d\tau}\, d\tau = \int_a^b ||v(t)||\, dt = \ell$$

lo que concluye la demostración. ▷

Finalizamos con el concepto de **ángulo** entre dos curvas.

Consideremos ahora dos curvas regulares suaves Γ_1 y Γ_2 tal que están parametrizadas por

$$\gamma_1 : [a, b] \to \mathbb{R}^3$$

$$\gamma_2 : [a, b] \to \mathbb{R}^3$$



respectivamente. Supóngase además que en las imágenes del punto t_0 se intersectan las curvas Γ_1 y Γ_2, esto es,

$$\gamma_1(t_0) = \gamma_2(t_0)$$

Si tales curvas están definidas en coordenadas x, y, z por

$$\begin{cases} \gamma_1(t) = (x_1(t),\, y_2(t),\, z_1(t)) \\ \gamma_2(t) = (x_2(t),\, y_2(t),\, z_2(t)) \end{cases}$$

y son tales que sus vectores tangentes en el punto común $\gamma_1(t_0) = \gamma_2(t_0)$ son

$$\xi = \left(\frac{dx_1}{dt}, \frac{dy_1}{dt}, \frac{dz_1}{dt}\right)|_{t=t_0},$$

$$\eta = \left(\frac{dx_2}{dt}, \frac{dy_2}{dt}, \frac{dz_2}{dt}\right)|_{t=t_0},$$

entonces es natural definir el **ángulo** formado por esas curvas en $\gamma_1(t_0) = \gamma_2(t_0)$ como el ángulo formado por sus vectores tangentes. Esto es, si tal ángulo es θ, entonces

$$\cos\theta = \frac{<\xi, \eta>}{||\xi||||\eta||}.$$

La figura 7.5 ilustra esta discusión.

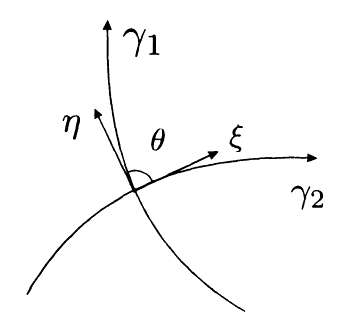


Figura 7.5: Ángulo entre dos curvas.

EJEMPLO. Sean las curvas Γ_1 y Γ_2 definidas por

$$\Gamma_1 : \gamma_1(t) = (t, t^2) \quad \text{con} \quad t \in [0, 4]$$

$$\Gamma_2 : \gamma_2(t) = (t, t^3) \quad \text{con} \quad t \in [0, 4]$$



◁ Es claro que $\gamma_1(1) = \gamma_2(1) = (1,1)$ es el punto de intersección de las curvas Γ_1 y Γ_2. Si nombramos a los vectores tangentes respectivos por

$$\xi = \dot{\gamma}_1(1) = (1,2)$$

$$n = \dot{\gamma}_2(1) = (1,3)$$

entonces el ángulo formado en $(1,1)$ por las curvas Γ_1 y Γ_2 es

$$\theta = \arccos \frac{<\xi, n>}{||\xi||\,||n||} = \arccos\left(\frac{<(1,2),(1,3)>}{\sqrt{5}\sqrt{10}}\right)$$

$$= \arccos\left(\frac{7}{\sqrt{50}}\right) = \arccos\left(\frac{7\sqrt{2}}{10}\right) \quad \triangleright$$

7.3 Ejercicios

1. Dé una orientación a las curvas definidas en los incisos **a, b, c** y **g** en el ejercicio **1.** del capítulo 4.

2. a. Calcule la longitud de arco de 0 a π en las curvas de los incisos del ejercicio 1.

b. Calcular el ángulo entre las curvas dadas en los puntos donde se intersectan.

$$\gamma_1(t) = (2\cos t,\ 2\,\mathrm{sen}\, t),\ \gamma_2(t) = (t,\ t^2)$$

$$\gamma_1(t) = (e^t\cos t,\ e^t\,\mathrm{sen}\, t),\ \gamma_2(t) = (t, t)$$

$$\gamma_1(t) = (t^2, t^3),\ \gamma_2(t) = (t^3, t^2).$$

Indicación: calcule primero los puntos de intersección.

Capítulo 8

Límites y puntos singulares

En este capítulo discutimos someramente el concepto de límite de una función real de variable vectorial en un **punto de acumulación** del dominio. Los temas son tratados de igual forma que en el capítulo 2 de **Reyes** (1996).

8.1 Puntos de acumulación y límites

Sea $\Omega \subset \mathbb{R}^n$ una región del espacio euclidiano, de dimensión n, donde consideramos apenas $n = 2, 3$.

DEFINICIÓN. El punto $p \in \mathbb{R}^3$ se dirá que es un **punto de acumulación** para la región Ω, sí toda bola $B_\epsilon(p)$ intersecta no vacíamente a Ω. Esto es, para cada bola de radio $\epsilon > 0$ con centro en p se encuentra un punto $q \in \Omega$ contenido en tal bola.

Si p no es de acumulación para Ω, se dirá un punto **aislado** (véase la figura 8.1).

En términos menos precisos, un punto de acumulación p de la región Ω es aquél que está tan próximo de Ω como se desée. El punto p no necesariamente está contenido en Ω, pero se puede aproximar a él por puntos contenidos en Ω.

EJEMPLO. Sea $\Omega = (a, b) \subset \mathbb{R}^1$, es claro que $p = a$ es un punto de acumulación de Ω. Todo punto $q \in (a, b)$ es de acumulación.

EJEMPLO. Sea $\Omega = \mathbb{R}^2 - \{(0,0)\}$, entonces $p(0,0)$ es un punto de acumulación de Ω. También todo punto $q \neq (0,0)$ contenido en Ω es de

acumulación para Ω.

EJEMPLO. Sea la región espacial

$$\Omega = \{(x, y, z) | \, x + y + z < 1\}$$

$\lhd$ Es claro que el punto frontera $p = (0, 0, 1)$ es un punto de acumulación de Ω. Por otro lado, el punto $q = (1, 1, 1)$ no pertenece a Ω y además es un punto aislado de Ω. $\rhd$

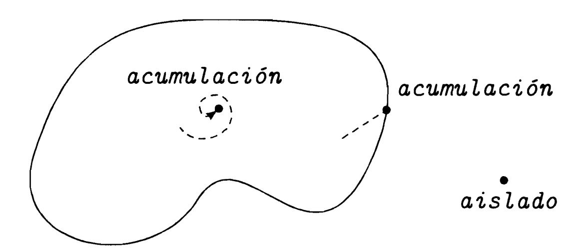


Figura 8.1: Puntos de acumulación y aislados.

Definamos ahora el concepto de **límite** de una función real de variable vectorial en un punto de acumulación de su dominio, en términos de los límites de funciones reales de variable real.

Sea $f : \Omega \to \mathbb{R}$ una función real de variable vectorial, y sea $p \in \mathbb{R}^3$ un punto de acumulación de Ω. Entendamos por $q = (x, y, z)$ el vector variable y entendemos que el punto $q \in \Omega$ se aproxima al punto p por la notación $\lim_{q \to p}$.

Además, si $p = (p_1, p_2, p_3)$ son las coordenadas, el número real $||q - p||$ se entiende como

$$||q - p|| = \sqrt{(x - p_1)^2 + (y - p_2)^2 + (z - p_3)^2}.$$

DEFINICIÓN. Se dice que la función $f(x, y, z)$ tiene límite L en el punto de acumulación p de Ω, definido por

$$\lim_{q \to p} f(x, y, z) = L$$

si y sólo sí, se cumple que existe el límite real, de variable real

$$\lim_{||q - p|| \to 0} (f(x, y, z) - L) = 0$$

Ya que conocemos del cálculo diferencial básico, el cálculo de los límites de variable real, así como su existencia, se tiene la siguiente lista de resultados para los límites en variables vectoriales.



LEMA 8.1 *El límite de la función real de variable vectorial,*

$$\lim_{q \to p} f(x, y, z) = L$$

cuando existe es único.

TEOREMA 8.1 (Aritmética de límites) *Sean* $f, g : \Omega \to \mathbb{R}$ *dos funciones reales de variable vectorial y* $p \in \mathbb{R}^3$ *un punto de acumulación de la región* Ω *tales que,*

$$\lim_{q \to p} f(x, y, z) = L_1, \quad \lim_{q \to p} g(x, y, z) = L_2$$

entonces

a. $\lim_{q \to p}(f \pm g)(x, y, z) = \lim_{q \to p} f(x, y, z) \pm \lim_{q \to p} g(x, y, z) = L_1 \pm L_2$

b. $\lim_{q \to p}(fg)(x, y, z) = \lim_{q \to p} f(x, y, z) \lim_{q \to p} g(x, y, z) = L_1 L_2$

c. *Cuando* $L_2 \neq 0$ *se tiene además que*

$$\lim_{q \to p} \left(\frac{f}{g}\right)(x, y, z) = \frac{\lim f(x, y, z)}{\lim g(x, y, z)} = \frac{L_1}{L_2}$$

Estos resultados nos garantizan el cálculo de límites de funciones en puntos de acumulación que están contenidos en el dominio. A continuación damos un par de ejemplos, donde se aplica apenas una simple sustitución.

EJEMPLO. Sea la función

$$f(x, y, z) = \ln(1 - x - y - z) + x^2 + y^2 + z^2 + 1$$

$\lhd$ Es claro que el punto $p = (0, 0, 0)$ está contenido en el dominio Ω de f y es un punto de acumulación de Ω. Por el teorema anterior se tiene que

$$\lim_{(x,y,z) \to (0,0,0)} f(x, y, z) = \lim_{(x,y,z) \to (0,0,0)} [\ln(1 - x - y - z) + x^2 + y^2 z^2 + 1] = 1$$

que se obtiene de una sustitución simple. $\rhd$

EJEMPLO. Calcular el límite,

$$\lim_{(x,y) \to (1,1)} e^{x+y}$$

$\lhd$ Una sustitución directa nos lleva a que

$$\lim_{(x,y) \to (1,1)} e^{x+y} = e^{1+1} = e^2 \quad \rhd$$

El problema del cálculo de límites de variable vectorial se complica cuando el punto de acumulación p no está en el dominio natural de la función.



8.2 Puntos singulares

Un caso interesante de un punto de acumulación es aquél cuando no está contenido en el dominio de una función por generar una indeterminación al momento de tratar de evaluarlo. Si en el intento de evaluación genera un cociente por cero se dice que tal punto es una **singularidad algebráica** de la función (véase el capítulo 2 de Reyes, 1996).

EJEMPLO. Consideremos nuevamente nuestra vieja conocida función

$$f(x, y) = \frac{xy}{x^2 + y^2}$$

analizada en el capítulo 4 y mostrada en la figura 4.13.

$\lhd$ Su dominio es la región $\Omega = \mathbb{R}^2 - \{(0,0)\}$, siendo el punto $p = (0,0)$ de acumulación de Ω.

a. Tomemos una aproximación en Ω a p a través el eje x, es decir mediante los puntos de la forma $(x, 0)$, y en este caso se tiene

$$\lim_{(x,y)\to p} f(x,y) = \lim_{(x,0)\to(0,0)} \frac{xy}{x^2+y^2} = \lim_{x\to 0} \frac{0}{x^2} = \lim_{x\to 0} 0 = 0$$

b. Consideremos ahora una aproximación en Ω a p a través del eje y, es decir, mediante los puntos de la forma $(0, y)$, y en este caso se tiene

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(0,y)\to(0,0)} \frac{xy}{x^2+y^2} = \lim_{y\to 0} \frac{0}{y^2} = \lim_{y\to 0} 0 = 0$$

c. Si hacemos la aproximación en Ω al punto p por la recta identidad $y = x$ se obtiene

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,x)\to(0,0)} \frac{xy}{x^2+y^2} = \lim_{x\to 0} \frac{x^2}{x^2+x^2} = \lim_{x\to 0} \frac{x^2}{2x^2} = \frac{1}{2}$$

De esta manera, hemos encontrado que existen trayectorias que se aproximan al punto $p = (0,0)$ y sus límites correspondientes no son iguales. Ya que el límite tiene que ser único por cualquier forma de aproximación, se concluye que el límite

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$$

no existe. $\rhd$

Una forma más general de estudiar un límite como el del ejemplo anterior es usar un **cambio de coordenadas** . Esto es, utilizar otras variables



que identifiquen una posible dependencia del límite en una dirección o de una curva.

EJEMPLO. Sea nuevamente el límite

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$$

$\lhd$ Usemos el siguiente cambio de **coordenadas polares**

$$\begin{cases} x = r\cos\theta \\ y = r\operatorname{sen}\theta \end{cases}$$

Entonces el cociente dado se expresa

$$\frac{xy}{x^2+y^2} = \frac{r^2\cos\theta\operatorname{sen}\theta}{r^2} = \cos\theta\operatorname{sen}\theta$$

De donde, en virtud de que $(x,y)\to(0,0) \iff r\to 0$, se tiene que

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} = \lim_{r\to 0}\cos\theta\operatorname{sen}\theta = \cos\theta\operatorname{sen}\theta$$

que depende del ángulo θ con el cual se aproximen las variables (x,y) al origen.

Por ejemplo, sobre la recta $y=0$ (eje x) el ángulo es $\theta=0$ lo que implica que $\operatorname{sen}\theta=0$ y, por lo tanto, el límite buscado sobre esa recta es cero.

Análogamente, sobre la recta $x=0$ (eje y) el ángulo es $\theta=\frac{\pi}{2}$, lo que implica que $\cos\theta=0$ y, por lo tanto, el límite buscado sobre tal recta es igualmente cero.

En el caso de calcular sobre la recta $y=x$, el ángulo es $\theta=\pi/4$ y, por lo tanto, $\cos\pi/4 \operatorname{sen}\pi/4 = \frac{\sqrt{2}}{2}.\frac{\sqrt{2}}{2} = \frac{2}{4} = 1/2$ que es el límite calculado anteriormente. $\rhd$

EJEMPLO. Calcular el límite

$$\lim_{(x,y)\to(0,0)} \frac{x^2+y^2}{\sqrt{x^2+y^2+1}-1}$$

$\lhd$ Al usar el cambio de coordenadas (**polares**)

$$\begin{cases} x = r\cos\theta \\ y = r\ \operatorname{sen}\ \theta \end{cases}$$



se tiene que $x^2 + y^2 = r^2$, de donde

$$\frac{x^2 + y^2}{\sqrt{x^2 + y^2 + 1} - 1} = \frac{r^2}{\sqrt{r^2 + 1} - 1}$$

Ya que $(x, y) \to (0,0) \iff r \to 0$ se tiene que,

$$\lim_{(x,y)\to(0,0)} \frac{x^2 + y^2}{\sqrt{x^2 + y^2 + 1} - 1} = \lim_{r\to 0} \frac{r^2}{\sqrt{r^2 + 1} - 1}$$

$$= \lim_{r\to 0} \frac{r^2\left(\sqrt{r^2+1}+1\right)}{\left(\sqrt{r^2+1}-1\right)\left(\sqrt{r^2+1}+1\right)} = \lim_{r\to 0} \frac{r^2\left(\sqrt{r^2+1}+1\right)}{(r^2+1)-1}$$

$$= \lim_{r\to 0} \frac{r^2\left(\sqrt{r^2+1}+1\right)}{r^2} = \lim_{r\to 0}\left(\sqrt{r^2+1}+1\right) = 2$$

Por lo tanto,

$$\lim_{(x,y)\to(0,0)} \frac{x^2 + y^2}{\sqrt{x^2 + y^2 + 1} - 1} = 2 \quad \triangleright$$

EJEMPLO. Calcular el límite

$$\lim_{(x,y)\to(0,0)} \frac{(x+y)^2}{x^2+y^2}$$

$\triangleleft$ Nuevamente, usando el cambio de coordenadas anterior,

$$x = r\cos\theta, \quad y = \operatorname{sen}\theta$$

se tiene que

$$\frac{(x+y)^2}{x^2+y^2} = \frac{(r\cos\theta + r\operatorname{sen}\theta)^2}{r^2} = \frac{r^2(\cos\theta + \operatorname{sen}\theta)^2}{r^2} = (\cos\theta + \operatorname{sen}\theta)^2$$

que depende de la variable θ.

Si se toma por ejemplo $\theta = 0$ (eje x) se tiene que,

$$\lim_{(x,y)\to(0,0)} \frac{(x+y)^2}{x^2+y^2} = \lim_{r\to 0}(\cos\theta + \operatorname{sen}\theta)^2 = (\cos 0 + \operatorname{sen} 0)^2 = 1$$

Por otro lado, si se toma $\theta = \pi/4$ (la recta $y = x$) se tiene que,

$$\lim_{(x,y)\to(0,0)} \frac{(x+y)}{x^2+y^2} = \lim_{r\to 0}(\cos\pi/4 + \operatorname{sen}\pi/4) = \left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\right)^2 = (\sqrt{2})^2 = 2,$$

lo que nos dice que tal límite no existe. $\triangleright$

8.3 Continuidad

Ahora extendemos el concepto de continuidad que conocemos para funciones reales de variable real, para funciones reales de variable vectorial.

DEFINICIÓN. Sea $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ una función y sea $p \in \Omega$ un punto arbitrario del dominio. Decimos que la función f es **continua en el punto** p si

a. p es un punto aislado de Ω.

b. cuando p es punto de acumulación de Ω, entonces

$$\lim_{(x,y,z)\to p} f(x,y,z) = f(p)$$

Si la función f es continua en cada uno de los puntos de su dominio se dirá simplemente que f es **continua**.

Del teorema sobre la aritmética de límites se tiene que la suma (y diferencia), el producto y el cociente de funciones continuas es una función continua.

Más aún, en todas las funciones vectoriales donde intervengan las funciones clásicas: polinomios, exponenciales, trigonométricos, etc, con varias variables independientes, se tendrán funciones continuas.

EJEMPLO. A continuación todas las funciones que se mencionan son continuas (redundante decir que en sus dominios naturales).

a. $f(x,y) = \frac{x}{x+y^2}$

b. $f(x,y) = \operatorname{arcsen}(x^2+y^2)$

c. $f(x,y,z) = \frac{x+y-z}{x^2+y^2-z^2}$

d. $f(x,y,z) = e^{x+y^2} \operatorname{sen}\left(\frac{1}{z}\right)$

e. $f(x,y) = \frac{8x^2y^2+z^2}{\sqrt{\ln z + \ln x}} + e^{\cos(xyz)}$

Como es de dominio general, un problema interesante es aquél cuando una función con una singularidad p precisa de ser extendida continuamente en tal punto. Un resultado sobre la extensión continua es el siguiente.

TEOREMA 8.2 *Si f es una función continua en la región $\Omega \subset \mathbb{R}^3$ y p es una singularidad de f, para que f tenga una extensión continua F en el dominio $D = \Omega \cup \{p\}$ es necesario y suficiente que*

$$\lim_{(x,y,z)\to p} f(x,y,z) = L$$

exista.



◁ En este caso, la extensión continua de f en p está dada por la función

$$F(x,y,z) = \begin{cases} f(x,y,z) & \text{si} \quad (x,y,z) \in \Omega \\ L & \text{si} \quad (x,y,z) = p \end{cases} \qquad \triangleright$$

A continuación ilustramos el teorema 8.2 con un ejemplo.

EJEMPLO. Considérese la función

$$f(x,y) = \frac{x^3y^2}{x^2+y^2}$$

¿Se puede extender continuamente la función en el punto singular $p = (0,0)$?

◁ Ciertamente el dominio de la función f es $\Omega = \mathbb{R}^2 - \{(0,0)\}$, siendo $p = (0,0)$ un punto singular de f.

De manera análoga al caso de una sola variable, calculamos el límite

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} \frac{x^3y^2}{x^2+y^2}$$

Usando el cambio de variables mencionado anteriormente,

$$\begin{cases} x = r\cos\theta \\ y = r\,\mathrm{sen}\,\theta \end{cases}$$

se tiene que

$$\frac{x^3y^2}{x^2+y^2} = \frac{r^3\cos^3\theta\, r^2\,\mathrm{sen}^2\theta}{r^2} = r^2\cos^3\ \mathrm{sen}^2\,\theta.$$

De esta manera,

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} \frac{x^3y^2}{x^2+y^2} = \lim_{r\to 0} r\cos^3\theta\ \mathrm{sen}^2\theta = 0$$

Por lo tanto, la extensión continua de la función f en el punto está dada por la nueva función,

$$F(x,y) = \begin{cases} \frac{x^3y^2}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases} \qquad \triangleright$$

Un resultado importante de la continuidad de una función que es de interés en las aplicaciones en el siguiente

TEOREMA 8.3 (del valor intermedio). *Si $f : \Omega \to \mathbb{R}$ es una función continua definida en una región (conexa) y para los puntos $p, q \in \Omega \subset \mathbb{R}^3$ se tiene que $f(p) < f(q)$, entonces para todo valor $c \in [f(p), f(q)]$ existe un punto $r \in \Omega$, tal que $f(r) = c$.*

$\lhd$ Sea $\gamma : [0,1] \to \Omega$ una curva continua tal que una a los puntos p y $q \in \Omega$, es decir, tal que $\gamma(0) = p$ y $\gamma(1) = q$.

Consideremos ahora la composición real de variable real,

$$g = f \circ \gamma : [0,1] \to \Omega \to \mathbb{R}$$

dada por

$$g(t) = (f \circ \gamma)(t) = f(\gamma(t))$$

Entonces, es fácil ver que g es una función real de variable real continua, tal que $g(0) = f(p), g(1) = f(q)$. Por el teorema del valor intermedio para funciones continuas reales de variable real, para $c \in [f(p), f(q)]$ existe un argumento $t_0 \in [0,1]$ tal que $g(t_0) = c$. Defínase $r = \gamma(t_0)$, entonces se tiene que $f(r) = f(\gamma(t_0)) = g(t_0) = c$, lo que termina la prueba. $\rhd$

Ejercicios

1. Calcular los siguientes límites.

a. $\lim_{(x,y)\to(1,2)}(4x + 6y^2)$

b. $\lim_{(x,y)\to(-1,1)} \frac{x}{\sqrt{y-x}}$

c. $\lim_{(x,y)\to(1,0)} = \frac{\text{arcsen}\,(y/x)}{1+xy}$

d. $\lim_{(x,y)\to(2,2)} \frac{xy}{y^2+x^2}$

e. $\lim_{(x,y)\to(0,0)} \frac{e^{xy}}{4+xy}$

2. Analizar las siguientes funciones en el punto singular situado en el origen. ¿En cuáles singularidades se puede extender continuamente la función?

a. $f_1(x,y) = \frac{y^4-x^4}{x^2+y^2}$

b. $f_2(x,y) = \frac{4x^2y}{x^2+y^2}$

c. $f_3(x,y) = \frac{3x^2y}{1+x^4+2y^2}$

d. $f_4(x,y) = \frac{\text{sen}\,(xy)}{x^2+y^2}$

e. $f_5(x,y) = \frac{y^2+xy-2x^2}{y-x}$

f. $f_6(x,y) = \frac{e^x-e^y}{e^{-x}-e^{-y}}$



g. $f_7(x, y) = (x + y)\text{sen}\,\frac{1}{x}$

h. $f_8(x, y) = (x + y)\,\text{sen}\,\frac{1}{x+y}$



Capítulo 9

Valores extremos de funciones $\mathbb{R}^2 \to \mathbb{R}$

En este capítulo hacemos el análisis local de una función de clase C^r alrededor de un punto p, para el caso en que la función está definida en una región Ω del plano $\mathbb{R}^2$. Los resultados aquí obtenidos se generalizan para más variables y damos ejemplos de estas generalizaciones.

9.1 Plano tangente

En esta sección introducimos el concepto de **espacio tangente** en un punto de un conjunto de nivel de una función real de variable vectorial, mediante el uso del vector gradiente.

Consideremos el caso $\mathbb{R}^n = \mathbb{R}^3$, y sean, $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ una función suave, c un valor arbitrario en $\mathbb{R}$ y la **superficie de nivel** c de f,

$$S = f^{-1}(c) = \{(x, y, z) \in \Omega |\, f(x, y, z) = c\}$$

Si $\gamma : (a, b) \to \Omega$ es una curva tal que $\gamma(t) \in S$, entonces para todo $t \in (a, b)$ se cumple $f(\gamma(t)) = c$.

$\triangleleft$ Al derivar cada lado de la igualdad respecto a t se tiene

$$\frac{d}{dt}(f \circ \gamma)(t) = 0$$

y por La regla de la cadena, en todo el intervalo (a, b), se cumple la igualdad

$$\langle \nabla f_{(\gamma(t))}, \dot{\gamma}(t) \rangle = 0$$

Tomemos un punto arbitrario $q \in S$ y una curva γ tal que $\gamma(0) = q$. Entonces $\dot{\gamma}(0)$ es un vector tangente a γ en q.

De esto, se tiene que

$$0 = \langle \nabla f_{\gamma(t)}, \dot{\gamma}(t) \rangle \, |_{t=0} = \langle \nabla f_q, \dot{\gamma}(0) \rangle$$

Concluimos entonces que el vector gradiente ∇f_q es ortogonal al vector tangente $\dot{\gamma}(0)$.

Resumiendo, dado un punto $q \in S$ y una curva $\gamma(t) \in S$ tal que $\gamma(0) = q$ se tiene que el gradiente de f en q es ortogonal al vector tangente $\dot{\gamma}(0) = v$. Como la curva $\gamma(t)$ pasando por el punto q es arbitraria, entonces se puede realizar la siguiente definición.

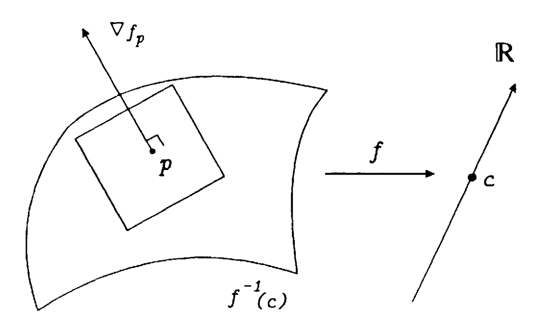


Figura 9.1: Plano tangente a una superficie de nivel.

DEFINICIÓN. Definimos el **plano tangente** T_qS a la superficie de nivel S en el punto q, como el plano pasando por q ortogonal al vector ∇f_q, y lo definimos por T_qS (véase la figura 9.1).

Hemos supuesto, en la definición dada, que necesariamente $\nabla f_q \neq (0, 0, 0)$.

EJEMPLO. Calcular el plano tangente a la esfera $x^2 + y^2 + z^2 = 3$ en el punto $q = (1, 1, 1)$.

◁ La relación

$$x^2 + y^2 + z^2 = 3$$

implica que

$$x^2 + y^2 + z^3 - 3 = 0$$

Sea $f : \mathbb{R}^3 \to \mathbb{R}$ la función definida por

$$f(x, y, z) = x^2 + y^2 + z^2 - 3$$

entonces la esfera dada se escribe como la superficie de nivel

$$f^{-1}(0) = \{(x, y, z)|\, f(x, y, z) = 0\} = \{(x, y, z)|\, x^2 + y^2 + z^2 - 3 = 0\}$$

El punto $(1, 1, 1)$ pertenece al conjunto de nivel $f^{-1}(0)$, pues $1^2 + 1^2 + 1^2 - 3 = 0$

Después, calculamos el vector gradiente $\nabla f_{(1,1,1)}$,

$$\nabla f_{(1,1,1)} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)|_{(1,1,1)} = (2x, 2y, 2z)|_{(1,1,1)} = (2, 2, 2)$$

Por lo tanto, la ecuación del plano tangente a $S = f^{-1}(0)$ en $(1, 1, 1)$ se calcula

$$< (2, 2, 2), (x, y, z) >=< (2, 2, 2), (1, 1, 1) >$$

es decir, la ecuación buscada es,

$$2x + 2y + 2z = 6 \quad \rhd$$

Mencionamos que la metodología utilizada para calcular planos tangentes a superficies se aplica también para calcular en general, los espacios tangentes de conjuntos de nivel de funciones suaves. Particularmente, el método se usa para calcular rectas tangentes de curvas planas, como lo muestra el siguiente ejemplo.

EJEMPLO. Calcular la recta tangente a la curva

$$C: \quad x^2y + y^3 = 10$$

en el punto $(1, 2)$.

$\lhd$ Consideremos la función $f(x, y) = x^2y + y^3$. Entonces, la curva dada C está definida como la curva de nivel,

$$f^{-1}(10) = \{(x, y)|\, f(x, y) = 10\} = \{(x, y)|\, x^2y + y^3 = 10\} = C.$$

Verificamos primeramente que el punto $(1, 2) \in f^{-1}(10)$, pues

$$1^2(2) + (2)^3 - 2 + 8 = 10$$

y después calculamos el gradiente

$$\nabla f_{(1,2)} = (2xy,\, x^2 + 3y^2)_{(1,2)} = (4, 13)$$

Por lo tanto, la ecuación del tangente (que es una recta) $T_{(2,2)}C$ es,

$$< (4, 13),\, (x, y) >=< (4, 13)(1, 2) >$$

o equivalentemente,

$$4x + 13y = 30 \quad \rhd$$

9.2 Puntos regulares y críticos

Sean, $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ una función de clase C^r y $p = (a, b)$ un punto en Ω. Entonces, el desarrollo de Taylor de f alrededor de p hasta orden 2 esta dado por

$$\begin{aligned} f(x, y) = f(p) + \frac{\partial f}{\partial x}(p)(x - a) + \frac{\partial f}{\partial y}(p)(y - b) \\ + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(p)(x - a)^2 + \frac{\partial^2 f}{\partial y \partial x}(p)(x - a)(y - b) \\ + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(p)(y - b)^2 + o((x - a)^3, (y - b)^3) \end{aligned}$$

si el punto variable (x, y) está próximo del punto p en Ω.

DEFINICIÓN. Si el punto p es tal que $\nabla f_p \neq (0, 0)$, entonces se dirá un **punto regular** de f, y $f(p)$ se dirá **un valor regular**.

$\lhd$ En este caso, podemos suponer que $\frac{\partial f}{\partial x}(p) \neq 0$, y por El **teorema de Taylor** , la función se escribe localmente como la aproximación

$$z = f(p) + \frac{\partial f}{\partial x}(p)(x - a) + \frac{\partial f}{\partial y}(p)(y - b)$$

es decir, si $A = f(p)$, $B = \frac{\partial f}{\partial x}(p)$, $C = \frac{\partial f}{\partial y}(p)$, entonces

$$z = A + B(x - a) + C(y - b)$$

es la ecuación del plano tangente de la gráfica de la función f en el punto $(p, f(p)) = (a, b, f(a, b))$.

En otras palabras, alrededor de un punto regular la función f está aproximada por la ecuación del plano tangente en una vecindad del punto p. La figura 9.2 ilustra esta aproximación. $\rhd$

Esta discusión generaliza el concepto de la recta tangente a la gráfica de una función real de variable real de los cursos de cálculo básico, al del plano tangente a la gráfica (superficie) de una función de dos variables con valores reales. Dicho en otras palabras, la ecuación del plano tangente aproxima a la función alrededor del punto en cuestión.

EJEMPLO. Calcular aproximadamente $f(1.1, 1.1)$ si

$$f(x, y) = \ln(xy)$$

$\lhd$ Aproximamos f alrededor del punto $p = (1, 1)$, calculando

$$\frac{\partial f}{\partial x} = \frac{y}{x}, \quad \frac{\partial f}{\partial y} = \frac{x}{y}$$



lo que nos dice

$$B = \frac{\partial f}{\partial x}(1,1) = 1, \quad C = \frac{\partial f}{\partial y}(1,1) = 1$$

Por lo tanto, la ecuación del plano tangente de la gráfica de f en el punto $(1,1,0)$ es

$$z = f(1,1) + B(x-1) + C(y-1) = (x-1) + (y-1).$$

que aproxima a $f(x,y) = \ln(xy)$ alrededor del punto $(1,1)$.

De esto, la aproximación buscada se calcula por

$$z(1.1, 1.1) = (1.1 - 1) + (1.1 - 1) = 0.1 + 0.1 = 0.2$$

Esto es,

$$\ln(1.1)(1.1) \approx 0.2 \quad \triangleright$$

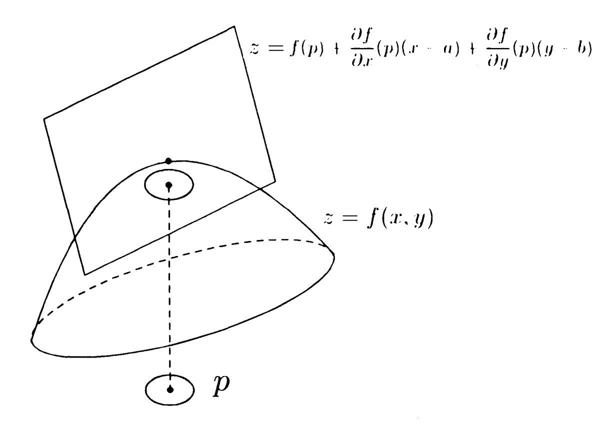


Figura 9.2: Aproximación lineal de f en p.

Cuando el punto p no es regular, el problema de aproximación de una función f se complica. Hacemos un análisis para el caso más simple de cuando no se tiene un punto regular.

DEFINICIÓN. Si $\nabla f_p = (0,0)$, es decir, las derivadas parciales de la función f se anulan en el punto p,

$$\frac{\partial f}{\partial x}(p) = 0 = \frac{\partial f}{\partial y}(p)$$

entonces p se dirá un **punto crítico** de f. Al valor $f(p)$ se le llamará un **valor crítico**.

Por El teorema de Taylor, alrededor del punto p la función f está aproximada por el polinomio cuadrático

$$z = f(p) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(p)(x-a)^2 + \frac{\partial^2 f}{\partial y \partial x}(p)(x-a)(y-b) + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(p)(y-b)^2$$

$$= f(p) + \frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(p)(x-a)^2 + 2\frac{\partial^2 f}{\partial y \partial x}(p)(x-a)(y-b) + \frac{\partial^2 f}{\partial y^2}(p)(y-b)^2\right]$$

lo cual puede escribirse matricialmente como

$$z = f(p) + \frac{1}{2}(x-a \quad y-b)\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(p) & \frac{\partial^2 f}{\partial y \partial x}(p) \\ \frac{\partial^2 f}{\partial x \partial y}(p) & \frac{\partial^2 f}{\partial y^2}(p) \end{pmatrix}\begin{pmatrix} x-a \\ y-b \end{pmatrix}$$

◁ Comprobamos que la multiplicación de las matrices señaladas en la última igualdad conducen a la expresión de la parte cuadrática de la aproximación. Al realizar el producto de las dos matrices finales, se obtiene la expresión

$$(x-a \quad y-b)\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(p)(x-a) + \frac{\partial^2 f}{\partial y \partial x}(p)(y-b) \\ \frac{\partial^2 f}{\partial y \partial x}(p)(x-a) + \frac{\partial^2 f}{\partial y^2}(p)(y-b) \end{pmatrix}$$

$$= \frac{\partial^2 f}{\partial x^2}(p)(x-a)^2 + \frac{\partial^2 f}{\partial y \partial x}(p)(y-b)(x-a) + \frac{\partial^2 f}{\partial y \partial x}(p)(x-a)(y-b) + \frac{\partial^2 f}{\partial y^2}(p)(y-b)^2$$

$$= \frac{\partial^2 f}{\partial x^2}(p)(x-a)^2 + 2\frac{\partial^2 f}{\partial y \partial x}(p)(x-a)(y-b) + \frac{\partial^2 f}{\partial y^2}(p)(y-b)^2$$

lo que sigue debido a la conmutatividad de las derivadas parciales mixtas de f en p. ▷

La matriz cuadrada obtenida de esta manera tiene un nombre y características especiales.

DEFINICIÓN. La matriz de 2×2 dada por

$$Hf_p = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(p) & \frac{\partial^2 f}{\partial y \partial x}(p) \\ \frac{\partial^2 f}{\partial y \partial x}(p) & \frac{\partial^2 f}{\partial y^2}(p) \end{pmatrix}$$

se llamará **matriz Hessiana** de la función f en el punto crítico p.



Es claro que La matriz Hessiana Hf_p es de la forma

$$A = \begin{pmatrix} M & L \\ L & N \end{pmatrix}$$

y, por lo tanto, es una matriz **simétrica**, es decir, coincide con su transpuesta,

$$A^t = \begin{pmatrix} M & L \\ L & N \end{pmatrix} = A$$

Tales matrices simétricas están estrechamente relacionadas con las llamadas **formas cuadráticas** que definen superficies cuadráticas.

Repasamos , entonces, un aspecto básico de las **formas cuadráticas** más simples en dos variables, que ya hemos estudiado en el capítulo 6.

Antes de iniciar la discusión de las formas cuadráticas, definimos los conceptos de valores máximo y mínimo local de una función de argumento vectorial.

DEFINICIÓN. Sean, $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ una función arbitraria y $p \in \Omega$ un punto en el dominio de f. Decimos que la función f tiene en $f(p)$ un

a. Valor **mínimo local**, si existe un número $\epsilon > 0$ tal que para cualquier punto (x, y) contenido en la bola de radio $\epsilon > 0$ y centro en p $(B_\epsilon(p))$, se tiene que $f(p) \leq f(x, y)$

b. Valor **máximo local**, si existe un número $\epsilon > 0$ tal que para cualquier $(x, y) \in B_\epsilon(p)$ se cumple que $f(x, y) \leq f(p)$

La figura 9.3 ilustra la definición.

Si para cualquier $(x, y) \in \Omega$ se cumple que $f(x, y) \geq f(p)$ entonces $f(p)$ se dirá un **mínimo global** de la función f en Ω. Análogamente se define un **máximo global** de la función f en Ω.

9.3 Formas cuadráticas básicas

DEFINICIÓN. Por una **forma cuadrática** de 2×2 entendemos a una matriz simétrica, real, de la forma

$$A = \begin{pmatrix} M & L \\ L & N \end{pmatrix}, \quad A^t = A.$$

tal que a una pareja de vectores $\xi = (\xi^1, \xi^2)$, $\eta = (\eta^1, \eta^2)$ los aplica en un número real $q(\xi, \eta)$ mediante la regla matricial

$$q(\xi, \eta) = (\eta^1, \eta^2) \begin{pmatrix} M & L \\ L & N \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix}$$

que definimos simplemente por $\eta^t A\xi$, siendo ξ un vector columna y η^t un vector renglón.

Así, una forma cuadrática aplicará en un sólo vector $\xi = (\xi^1, \xi^2)$ mediante la fórmula

$$q(\xi) = q(\xi, \xi) = \xi^t A\xi = (\xi^1 \ \xi^2) \begin{pmatrix} M & L \\ L & N \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix}$$
$$= M(\xi^1)^2 + 2L\xi^1\xi^2 + N(\xi^2)^2$$

generando una función $q : \mathbb{R}^2 \to \mathbb{R}$ que está caracterizada por la matriz dada A.

EJEMPLO. Sea la función cuadrática en las variables u, v dada por la expresión

$$q_1(u, v) = u^2 + v^2$$

$\lhd$ Entonces, escribiendo al vector $\xi = (u, v)$, podemos escribir matricialmente tal expresión como

$$q_1(u, v) = u^2 + 2(0)uv + v^2 = (u \ v) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

de donde q_1 tiene asociada la matriz simétrica (**forma cuadrática**)

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Por otro lado, en $p = (0, 0)$ la función tiene un **valor mínimo local** en $q_1(0, 0) = 0$

Esto se puede apreciar observando que la cuadrática $z = u^2 + v^2$ corresponde a un paraboloide circular que se abre verticalmente en dirección positiva del eje z. Si consideramos valores $c \in \mathbb{R}$ próximos de 0, entonces las imágenes inversas bajo f (curvas de nivel) $f^{-1}(c)$ corresponden a curvas cerradas que de hecho son círculos con ecuaciones $u^2 + v^2 = c$.

La figura 9.4 **a.** ilustra a la gráfica de esta forma cuadrática en una proximidad del origen de coordendas. $\rhd$

EJEMPLO. Ahora, consideremos la forma cuadrática dada por,

$$q_2(u, v) = -u^2 - v^2$$

$\lhd$ Esta tiene un **valor máximo local** en $f(0, 0) = 0$ y tiene asociada la matriz simétrica

$$A_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$



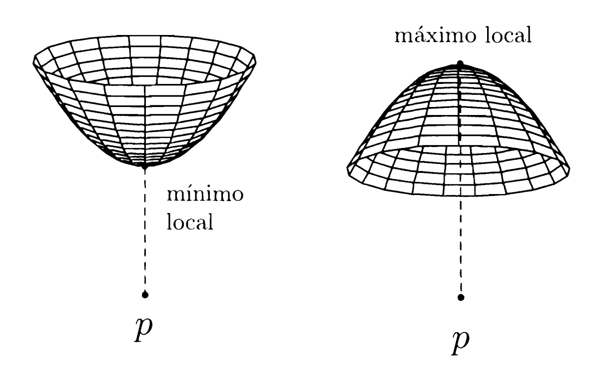


Figura 9.3: Valores máximos y mínimos locales.

De igual forma que para el caso anterior, se observa que la gráfica de $q_2(u, v)$ es un **paraboloide circular** que se abre verticalmente en dirección negativa del eje z. También, si consideramos valores negativos $c \in \mathbb{R}$ próximos del valor máximo $f(0) = 0$, las curvas de nivel correspondientes $u^2 + v^2 = -c$ son curvas cerradas, que además son círculos.

La figura 9.4 **b.** ilustra la gráfica de $q_2(u, v)$ es una proximidad del origen. $\triangleright$

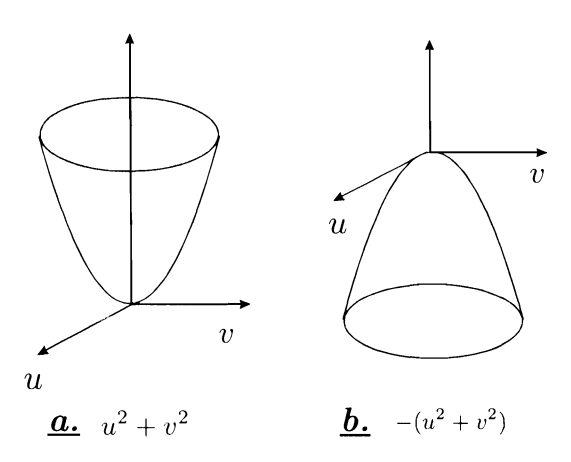


Figura 9.4: Formas cuadráticas $\pm(u^2 + v^2)$.

EJEMPLO. Consideremos la función cuadrática en las variables u, v dada por

$$q_3(u, v) = u^2 - v^2$$



que tiene asociada la matriz simétrica

$$A_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

debido a que, si ponemos el vector $\xi = (u, v)$ como un vector columna, matricialmente podríamos escribir

$$q_3(u, v) = u^2 - v^2 = (u\ v) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \xi^t A_3 \xi$$

Por otro lado, la cuadrática $q_2(u, v) = u^2 - v^2$ corresponde a la gráfica de la superficie **paraboloide hiperbólico**. Por esto, al tomar cualquier valor $c \in \mathbb{R}$ próximo del valor $f(0, 0) = 0$, la curva de nivel asociada $u^2 - v^2 = c$ corresponde a una hipérbola, mientras que la curva de nivel $f^{-1}(0)$ esta formado por la pareja de rectas $y = \pm x$ que se intersecan en el origen de coordenadas.

Por lo tanto, en $p = (0, 0)$ se tiene un **punto silla** $f(0, 0) = 0$ para la función q_3 (ni valor máximo, ni valor mínimo).

La figura 9.5 ilustra esta forma cuadrática. ▷

EJEMPLO. Análogamente, la forma cuadrática

$$q_4(u, v) = -u^2 + v^2$$

tiene asociada la matriz simétrica

$$A_4 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

y en el **punto silla** $p = (0, 0)$ tiene asociado un valor $f(0, 0) = 0$ llamado también **silla**.

Zoológico de formas cuadráticas básicas

Extendemos ahora el tratado general de una forma cuadrática, bosquejando a todas las formas cuadráticas básicas.

Dada la forma cuadrática

$$q(u, v) = au^2 + bv^2,$$

con a, b constantes diferentes de cero, esta función tiene asociada la matriz simétrica

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

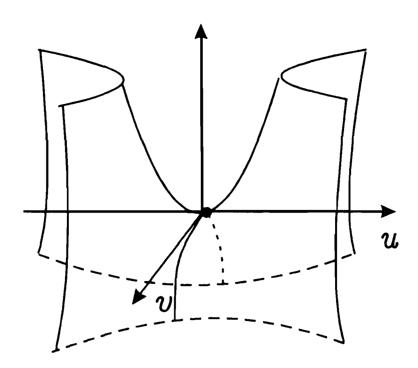


Figura 9.5: Forma cuadrática $u^2 - v^2$.

cuya gráfica tiene un comportamiento alrededor del punto $(p, f(p)) = (0, 0, 0)$ tiene alguna de las formas mostradas en la figura 9.6.

Observamos que las **formas cuadráticas** básicas no tienen términos cruzados en xy. Para cuando hay tales términos, estos se pueden omitir mediante el uso del siguiente resultado del **Algebra Lineal** y cuya prueba puede verse en Anton (1981).

LEMA 9.1 (de Sylvester de diagonalización) *Sea la matriz simétrica*

$$B = \begin{pmatrix} M & L \\ L & N \end{pmatrix}$$

Entonces, bajo un cambio lineal de coordenadas en el plano puede ser llevada a una **matriz diagonal** *de la forma*

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

donde λ_1 λ_2 son los **valores propios** *de A, que satisfacen el polinomio característico,*

$$0 = det\left[\begin{pmatrix} M & L \\ L & N \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]$$

Más aún se tiene que,

$$det\begin{pmatrix} M & L \\ L & N \end{pmatrix} = det\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \lambda_1\lambda_2$$



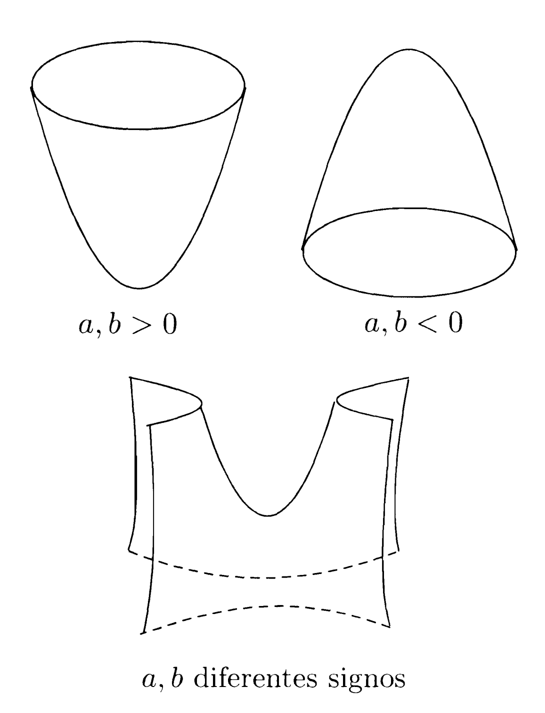


Figura 9.6: Zoológico de formas cuadráticas básicas.

Para ilustrar el alcance del Lema de Sylvester damos el siguiente ejemplo.

EJEMPLO. Sea la matriz simétrica

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

con determinante $det(B) = -3$.

$\triangleleft$ Procedemos a diagonalizar a la matriz B, calculando sus valores propios al resolver el polinomio cuadrático,

$$0 = det\left[\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] = det\begin{pmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 4$$

$$= \lambda^2 - 2\lambda - 3.$$

cuyas raíces son $\lambda_1 = 1, \ \lambda_2 = -3$.

De esta manera, la matriz diagonalizada es

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix}$$



cuyo determinante es $det(A) = -3$.

Por otro lado, la matriz A tiene asociada la forma cuadrática

$$q(u,v) = u^2 - 3v^2$$

que corresponde a un paraboloide hiperbólico (silla de montar) con un punto silla en el origen $(0,0,0)$.

Por lo tanto, la forma cuadrática asociada a la matriz B está dada en los vectores $\xi = (x,y)$ y matricialmente por,

$$\xi^T B \xi = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x^2 + 2(2xy) + y^2 = x^2 + 4xy + y^2$$

corresponde cualitativamente a un paraboloide hiperbólico con un punto silla en el origen $(0,0,0)$. $\triangleright$

El ejemplo anterior nos muestra que para analizar y caracterizar a un punto crítico p de una función $f(x,y)$ es suficiente con calcular su matriz Hessiana $B = Hf_p$ que es una matriz simétrica, y después diagonalizarla utilizando el Lema de Sylvester. Al obtener sus valores propios, conformamos la matriz diagonal

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

que caracteriza a una forma cuadrática básica de las mencionadas, para el caso en que los valores propios sean ambos diferentes de cero.

9.4 Puntos críticos no degenerados

DEFINICIÓN. Sea $f : \Omega \to \mathbb{R}$ una función de clase C^2 tal que en el punto crítico $p \in \Omega$ se tenga que el determinante de La matriz Hessiana

$$det(Hf_p) = det \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(p) & \frac{\partial^2 f}{\partial y \partial x}(p) \\ \frac{\partial^2 f}{\partial y \partial x}(p) & \frac{\partial^2 f}{\partial x^2}(p) \end{pmatrix} \neq 0$$

entonces llamamos al punto p un **punto crítico no degenerado**.

Observamos que si la matriz diagonalizada de Hf_p es la matriz

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

entonces será suficiente que ambos valores propios λ_1 y λ_2 sean diferentes de cero para que p sea un punto crítico no degenerado, en virtud de que por el Lema de Sylvester, $det(A) = det(Hf_p) = \lambda_1\lambda_2$.

EJEMPLO. Sea la función cuadrática en dos variables

$$f(x,y) = 3x^2 - 4xy + y^2$$

Analizar el comportamiento en sus puntos críticos mediante las matrices Hessianas correspondientes.

$\lhd$ El conjunto $\Omega = Dom(f) = \mathbb{R}^2$ es un dominio abierto y conexo. Realizaremos el análisis de los puntos críticos. Hallamos primeramente el gradiente, calculando las primeras derivadas parciales.

$$\frac{\partial f}{\partial x} = 6x - 4y$$
$$\frac{\partial f}{\partial y} = -4x + 12y$$

Es claro que

$$\nabla f_p = (0,0) \quad \text{si y sólo si,} \quad \left\{ \begin{array}{c} 6x - 4y = 0 \\ \text{y} \\ -4x + 2y = 0 \end{array} \right.$$

es decir, los puntos críticos se obtienen al resolver el sistema lineal

$$\left\{ \begin{array}{r} 6x - 4y = 0 \\ -4x + 2y = 0 \end{array} \right.$$

Como el determinante principal del sistema

$$D = \left| \begin{array}{cc} 6 & -4 \\ -4 & 2 \end{array} \right| = 12 - 16 = -4 \neq 0,$$

entonces, $x = 0$, $y = 0$ es la única solución de tal sistema. Esto es, $p = (0,0)$ es el único punto crítico, con valor cítico

$$f(0,0) = 3(0)^2 - 4(0)(0) + 0^2 = 0$$

Ahora calculamos las derivadas parciales segundas,

$$\frac{\partial^2 f}{\partial x^2} = 6, \quad \frac{\partial^2 f}{\partial y \partial x} = -4, \quad \frac{\partial^2 f}{\partial y^2} = 2$$



y obtenemos en el punto crítico $p(0,0)$,

$$\frac{\partial^2 f}{\partial x^2}(0,0) = 6, \quad \frac{\partial^2 f}{\partial y \partial x}(0,0) = -4, \quad \frac{\partial^2 f}{\partial y^2}(0,0) = 2$$

Por lo tanto la matriz Hessiana de f en el punto crítico $p(0,0)$ se calcula

$$Hf_{(0,0)} = \begin{pmatrix} 6 & -4 \\ -4 & 2 \end{pmatrix}$$

Despues, encontramos los valores propios λ_1, λ_2 para diagonalizar la matriz:

$$0 = \det \left[\begin{pmatrix} 6 & -4 \\ -4 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = \det \begin{pmatrix} 6-\lambda & -4 \\ -4 & 2-\lambda \end{pmatrix}$$

$$= \begin{vmatrix} 6-\lambda & -4 \\ -4 & 2-\lambda \end{vmatrix} = (6-\lambda)(2-\lambda) - 16 = \lambda^2 - 8\lambda - 4 = 0$$

con lo que,

$$\lambda = \frac{8 \pm \sqrt{64+16}}{2} = \frac{8 \pm \sqrt{80}}{20}$$

de donde,

$$\lambda_1 = \frac{8-\sqrt{80}}{2} < 0, \quad \lambda_2 = \frac{8+\sqrt{80}}{2} > 0$$

lo que implica que en el punto $(0,0,f(0,0))$ sobre la gráfica de f se tiene un punto silla, y además que el punto crítico $p = (0,0)$ es no degenerado.

Analizamos algunas **curvas de nivel** de f utilizando los resultados de la descripción de una forma cuadrática alrededor del origen de cordenadas.

a. Si se toma el valor $z = 1$ entonces

$$1 = 3x^2 - 4xy + y^2$$

es la ecuación de una hipérbola en el plano $\mathbb{R}^2$

b. Al tomar el valor $-1 = z$ entonces se tiene,

$$-1 = 3x^2 - 4xy + y^2$$

que es la ecuación de otra hipérbola en $\mathbb{R}^2$.

c. Si se considera $z = 0$ entonces se tiene la ecuación

$$0 = 3x^2 - 4xy + y^2$$



y para resolver en términos de la variable x nos da,

$$x = \frac{4y \pm \sqrt{16y^2 - 12y^2}}{6} = \frac{4y \pm 2y}{6} = \begin{cases} y \\ \frac{1}{3}y \end{cases}$$

es decir, la pareja de rectas

$$\begin{cases} y = x \\ y = 3x \end{cases}$$

que son las asíntotas de las hipérbolas dadas por la familia de ecuaciones

$$3x^2 - 4xy + y^2 = c$$

donde $c \in R$ es valor arbitrario de la función f.

La figura 9.7 describe la complejidad de la gráfica de la función f, así como sus curvas de nivel. $\triangleright$

Utilizamos el ejemplo anterior para hacer la siguiente observación.

$\triangleleft$ La función cuadrática inicial tenía la ecuación

$$f(x, y) = 3x^2 - 4xy + y^2$$

que matricialmente se puede escribir

$$3x^2 - 4xy + y^2 = 3x^2 + 2(-2)xy + y^2 = (x \quad y)\begin{pmatrix} 3 & -2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

siendo la matriz simétrica asociada (forma cuadrática)

$$A = \begin{pmatrix} 3 & -2 \\ -2 & 1 \end{pmatrix}$$

Por otro lado, encontramos que la matriz Hessiana en el punto crítico era la forma cuadrática

$$Hf_{(0,0)} = \begin{pmatrix} 6 & -4 \\ -4 & 2 \end{pmatrix} = 2A$$

Esto ocurre debido a que en la construcción de la matriz Hessiana durante la discusión general se factorizó el escalar $\frac{1}{2}$ dando paso a la Hessiana como se definió. $\triangleright$

Utilizamos ahora toda la herramienta que se ha desarrollado, para analizar puntos críticos de funciones suaves de dos variables.



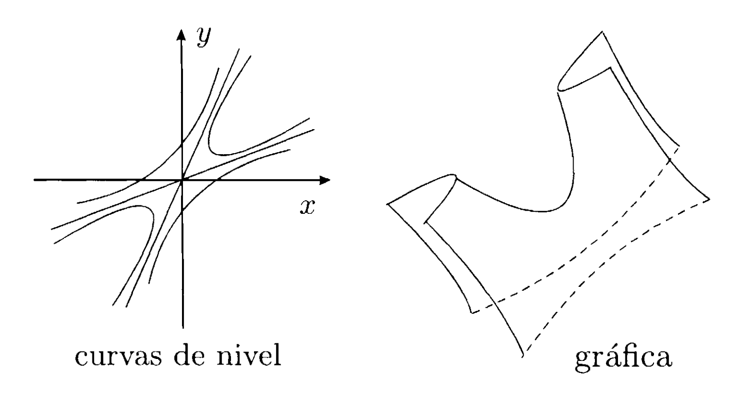


Figura 9.7: Curvas de nivel y gráfica de $f(x, y) = 3x^2 - 4xy + y^2$.

EJEMPLO. Sea la función polinomial de grado 4 en dos variables dada por la igualdad,

$$f(x, y) = x^4 + y^4 - 4xy$$

y cuyo dominio es $\Omega = \mathbb{R}^2$.

$\lhd$ Calculamos sus puntos críticos, igualando las coordenadas del gradiente a cero,

$$\begin{cases} \frac{\partial f}{\partial x} = 4x^3 - 4y = 0 \\ \frac{\partial f}{\partial y} = 4y^3 - 4x = 0 \end{cases}$$

o equivalentemente, resolviendo el sistema de ecuaciones

$$\begin{cases} x^3 - y = 0 \\ y^3 - x = 0 \end{cases}$$

Al despejar la variable y de la primera ecuación se tendrá

$$\begin{cases} y = x^3 \\ y^3 - x = 0 \end{cases}$$

y si se sustituye x^3 en la segunda ecuación se tendría $(x^3)^3 - x = 0$, lo que implicaría que $x^9 - x = 0$. Por lo tanto, es necesario resolver la ecuación de grado 9,

$$x(x^8 - 1) = 0$$

Al factorizar el lado izquierdo de esta ecuación se tiene que

$$x(x^4 - 1)(x^4 + 1) = 0 \Longleftrightarrow x(x^2 + 1)(x^2 - 1) = 0$$

lo que nos da las raíces $x = 0,\ 1, -1$.

Al sustituir tales raíces en la ecuación $y = \sqrt[3]{x}$ nos daría que $y = 0, 1, -1$.

Por lo tanto, los puntos críticos son $p_1 = (0,0)$, $p_2 = (1,1)$, $p_3 = (-1,-1)$ y sus valores críticos correspondientes serán,

$$f(0,0) = 0, \quad f(1,1) = -2, \quad f(-1,-1) = -2$$

Al calcular las derivadas parciales segundas de la función f se tiene

$$\frac{\partial^2 f}{\partial x^2} = 12x^2, \quad \frac{\partial^2 f}{\partial y \partial x} = -4, \quad \frac{\partial^2 f}{\partial y^2} = 12y^2$$

lo que da la matriz Hessiana en cada punto crítico p,

$$Hf_p = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}_p$$

Procedemos a analizar cada uno de los puntos críticos obtenidos, mediante la diagonalización de su correspondiente matriz Hessiana

a. Para el punto crítico $(0,0)$, su matriz Hessiana es,

$$Hf_{(0,0)} = \begin{pmatrix} 0 & -4 \\ -4 & 0 \end{pmatrix}$$

Para calcular los valores propios de esta matriz resolvemos la ecuación cuadrática característica.

$$0 = \begin{vmatrix} -\lambda & -4 \\ -4 & -\lambda \end{vmatrix} = \lambda^2 - 16$$

lo que implica que $\lambda = \pm 4$.

De esta forma, la matriz diagonal es la forma cuadrática

$$A_1 = \begin{pmatrix} 4 & 0 \\ 0 & -4 \end{pmatrix}$$

lo que nos dice que en el punto $(0,0,\ f(0,0)) = (0,0,0)$ hay un **punto silla**.

Además en una localidad del punto crítico $(0,0)$ las curvas de nivel son hipérbolas cuyas **separatrices** son dos curvas que se intersectan en el punto $(0,0)$, que es no degenerado.

b. En el punto crítico $(1,1)$, se tiene la matriz Hessiana

$$Hf_{(1,1)} = \begin{pmatrix} 12 & -4 \\ -4 & 12 \end{pmatrix}$$



Para calcular los valores propios resolvemos la ecuación cuadrática

$$0 = \begin{vmatrix} 12-\lambda & -4 \\ -4 & 12-\lambda \end{vmatrix} = (12-\lambda)^2 - 16 = \lambda^2 - 24\lambda + 128$$

esto es, $\lambda_1 = 16$, $\lambda_2 = 8$.

Lo anterior implica que $f(1,1) = -2$ es un mínimo local pues la matriz diagonalizada es, en este caso,

$$A_2 = \begin{pmatrix} 16 & 0 \\ 0 & 8 \end{pmatrix}$$

En virtud de que $f(1,1) = -2$ es un mínimo local, se tiene que las curvas de nivel de valores c próximos de -2 (de hecho mayores) son curvas cerradas que encierran al punto crítico no degenerado $(1,1)$

c. Análogamente al inciso anterior, en virtud de que en el punto crítico $(-1,1)$ la matriz Hessiana es,

$$Hf_{(-1,1)} = \begin{pmatrix} 12 & -4 \\ -4 & 12 \end{pmatrix}$$

la función tiene un mínimo local en $f(-1,-1) = -2$.

Una inspeccción simple de la gráfica de la función mediante los puntos críticos y sus valores críticos nos asegura la forma de las curvas de nivel de la función, las cuales vienen dadas por la familia de ecuaciones

$$x^4 + y^4 - 4xy = c$$

siendo $c \in R$ un valor que está condicionado a ser ≥ -2 que es el valor mínimo de la función.

Cuando $c = -2$ las curvas de nivel se reducen a la pareja de puntos críticos

$$p_1 = (1,1) \quad \text{y} \quad p_2 = (-1,1).$$

Si $-2 < c < 0$ entonces la ecuación $x^4+y^4-4xy = c$ describe una pareja de curvas disconexas que son cerradas y encierran a los puntos críticos p_1 y p_2.

Si $c = 0$, la ecuación $x^4 + y^4 - 4xy = 0$ describe una curva que es la unión de otras dos que se intersectan en el punto $(0,0)$ y conforman una figura en forma de "ocho".

Cuando $c > 0$ la curva de nivel asociada se vuelve un circuito que encierra a la figura de "ocho"



Ya anteriormente habíamos mostrado este ejemplo, y ahora con el uso de la Teoría diferenciable se ha podido hacer un análisis más preciso, pero que aún no está totalmente completado. La figura 9.8 ilustra cualitativamente el mapa topográfico de las curvas de nivel de la gráfica de la función. ▷

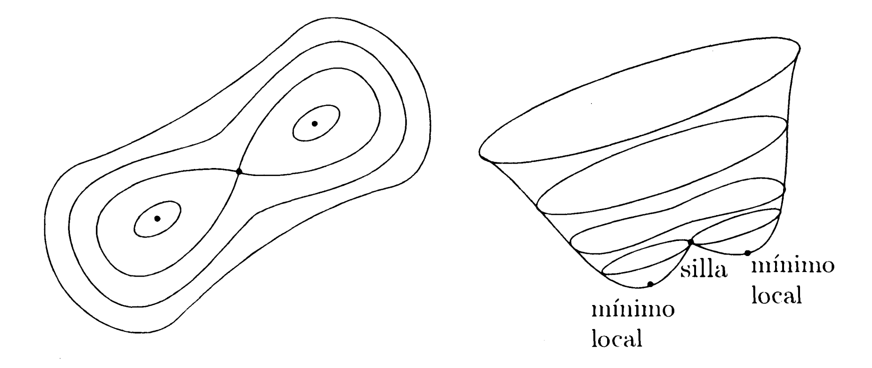


Figura 9.8: Mapa topográfico de $f(x, y) = x^4 + y^4 - 4xy$.

De hecho, aún para funciones polinomiales de dos variables, el análisis de una curva de nivel es complicado y requiere de metodología matemática moderna provista por áreas como la **Geometría Algebraica**. Hablando con la verdad, estas herramientas están muy fuera de nuestro alcance y apelamos a las construcciones de las gráficas de funciones de dos variables mediante las computadoras. Este análisis, aunque es cuantitativo y poco analítico, es de gran ayuda en los problemas prácticos.

Estos ejemplos muestran cuán complejos pueden ser las curvas de nivel asociadas a una función de dos variables.

Es claro que el problema se complica para dimensiones mayores. Así que en los subsecuentes ejemplos analizaremos apenas la forma de los puntos críticos, teniendo en cuenta la forma local de la gráfica en una vecindad de cada uno de ellos, cuando estos son no degenerados.

EJEMPLO. Analizar los puntos críticos de la función

$$f(x, y) = x^2 + x^2y + y^2 + 4$$

◁ Como es sabido, los puntos críticos se calculan al resolver la ecuación

$$\nabla f_p = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \right) = (0, 0)$$



es decir, al resolver el sistema de ecuaciones

$$\begin{cases} \frac{\partial f}{\partial x} = 2x + 2xy = 0 \\ \frac{\partial f}{\partial y} = 2y + x^2 = 0 \end{cases}$$

lo que equivale a que

$$\begin{cases} 2(x + xy) = 0 \\ 2y + x^2 = 0 \end{cases} \iff \begin{cases} x(1 - y) = 0 \\ y = -x^2/2 \end{cases} \iff \begin{cases} x = 0 \quad \text{ó} \quad y = -1 \\ \text{y} \\ y = \frac{-x^2}{2} \end{cases}$$

de donde

$$\begin{cases} x = 0 \quad \text{y} \quad y = \frac{-x^2}{2} \\ y = -1 \quad \text{y} \quad y = \frac{-x^2}{2} \end{cases} \iff \begin{cases} x = 0, \quad y = 0 \\ y = -1, \quad x = \pm\sqrt{2} \end{cases}$$

lo que implica que los puntos críticos buscados son

$$p_1 = (0, 0), \quad p_2 = (-\sqrt{2}, -1), \quad p_3 = (\sqrt{2}, -1)$$

Calculamos sus valores críticos correspondientes,

$$f(0, 0) = 4, \quad f(-\sqrt{2}, -1) = 5, \quad f(\sqrt{2}, -1) = 5$$

La matriz Hessiana de f en cualquier punto crítico p es

$$Hf_{(0,0)} = \begin{pmatrix} 2 + 2y & 2x \\ 2x & 2 \end{pmatrix}_p$$

y pasamos ahora a analizar cada punto crítico obtenido.

a. En el punto crítico $p_1 = (0, 0)$, tenemos la matriz Hessiana

$$Hf_{(0,0)} = \begin{pmatrix} 2 + 2y & 2x \\ 2x & 2 \end{pmatrix}_{(0,0)} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

que de hecho ya es una matriz diagonal.

No obstante, para afirmar que su diagonalización es la misma, procedemos a realizar tal diagonalización.

Para calcular los valores propios de la matriz $Hf_{(0,0)}$ se tiene que resolver la ecuación

$$0 = \begin{vmatrix} 2 - \lambda & 0 \\ 0 & 2 - \lambda \end{vmatrix} = (2 - \lambda)^2$$

lo que implica que $\lambda_1 = 2 = \lambda_2 > 0$, y por lo tanto, $f(0,0) = 4$ es un valor mínimo local.

b. Para el punto crítico $p_2 = (-\sqrt{2}, -1)$. la matriz Hessiana es, en este caso,

$$Hf_{(-\sqrt{2},-1)} = \begin{pmatrix} 0 & 2(-\sqrt{2}) \\ 2(-\sqrt{2}) & 2 \end{pmatrix} = \begin{pmatrix} 0 & -2\sqrt{2} \\ -2\sqrt{2} & 2 \end{pmatrix}$$

Calculamos los valores propios,

$$0 = \begin{vmatrix} 0-\lambda & -2\sqrt{2} \\ -2\sqrt{2} & 2-\lambda \end{vmatrix} = \begin{vmatrix} \lambda & -2\sqrt{2} \\ -2\sqrt{2} & 2-\lambda \end{vmatrix} = \lambda^2 - 2\lambda - 8$$

de donde, $\lambda_1 = 4$ $\lambda_2 = -2$. Esto es, la diagonal asociada es

$$A_2 = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$$

lo que nos dice que $f(-\sqrt{2}, -1) = 5$ corresponde a un valor crítico silla.

c. El punto crítico $p_3 = (\sqrt{2}, -1)$ tiene como matriz Hessiana a

$$Hf_{(\sqrt{2},-1)} = \begin{pmatrix} 0 & 2(\sqrt{2}) \\ 2(\sqrt{2}) & 2 \end{pmatrix}$$

Para calcular sus valores propios, es preciso resolver la ecuación

$$0 = \begin{vmatrix} 0-\lambda & 2\sqrt{2} \\ 2\sqrt{2} & 2-\lambda \end{vmatrix} = \begin{vmatrix} \lambda & 2\sqrt{2} \\ 2\sqrt{2} & 2-\lambda \end{vmatrix} = \lambda^2 - 2\lambda - 8$$

y obtenemos nuevamente, $\lambda_1 = 4$, $\lambda_2 = -2$. Por lo tanto, $f(\sqrt{2}, -1) = 5$ corresponde también a un valor silla.

Observamos que todos los puntos críticos de esta función son no degenerados, ya que todos los valores propios obtenidos no se anulan. ▷

EJEMPLO. Dada la función

$$f(x,y) = xy(1-x-y) = xy - x^2y - xy^2$$

analizar sus puntos críticos mediante el método mostrado.

◁ El gradiente de la función es

$$\nabla f_p = (y - 2xy - y^2, \quad x - x^2 - 2xy)_p$$

y se anula, si y sólo si, se satisface el sistema

$$\begin{cases} y - 2xy - y^2 = 0 \\ x - x^2 - 2xy = 0 \end{cases}$$

o equivalentemente,

$$\begin{cases} y(1-2x-y)=0 \\ x(1-x-2y)=0 \end{cases}$$

esto es,

$$\begin{cases} y=0 \quad \text{ó} \quad 1-2x-y=0 \\ \text{y} \\ x=0 \quad \text{ó} \quad 1-x-2y=0 \end{cases}$$

lo que implica que,

$$\begin{cases} y=0 \quad \text{y} \quad y=0 \\ x=0 \quad \text{y} \quad 1-2x-y=0 \\ y=0 \quad \text{y} \quad 1-x-2y=0 \\ 1-2x-y=0 \quad \text{y} \quad 1-x-2y=0 \end{cases}$$

de donde se obtienen los puntos críticos,

$$p_1=(0,0), \quad p_2=(0,1), \quad p_3=(1,0), \quad p_4=\left(\frac{1}{3},\frac{1}{3}\right)$$

con los valores críticos correspondientes

$$f(0,0)=0, \quad f(0,1)=0, \quad f(1,0)=0, \quad f(1/3,1/3)=1/27$$

Procedemos a analizar los puntos críticos sabiendo que la matriz Hessiana en cada uno de ellos viene dada por

$$Hf_p=\begin{pmatrix} -2y & 1-2x-2y \\ 1-2x-2y & 2x \end{pmatrix}$$

a. Para el punto crítico $p_1=(0,0)$, la matriz Hessiana está dada por

$$Hf_{(0,0)}=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

y sus valores propios se calculan resolviendo la ecuación

$$0=\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix}=\lambda^2-1$$

de donde $\lambda=\pm 1$. Esto dice que $f(0,0)=0$ corresponde a un valor silla.

b. En el punto $p_2=(0,1)$, la matriz Hessiana correspondiente es,

$$Hf_{(0,1)}=\begin{pmatrix} -2 & -1 \\ -1 & 0 \end{pmatrix}$$



y sus valores propios se calculan al resolver la ecuación,

$$0 = \begin{vmatrix} -2-\lambda & -1 \\ -1 & -\lambda \end{vmatrix} = -\lambda(-2-\lambda) - 1 = \lambda^2 + 2\lambda - 1$$

Encontramos los valores propios λ_1, λ_2 donde,

$$\lambda_1 = \frac{-2+\sqrt{8}}{2} > 0, \quad \lambda_2 = \frac{-2-\sqrt{8}}{2} < 0$$

y por lo tanto $f(0,1) = 0$ es un valor crítico silla.

c. El punto crítico $p_3 = (1,0)$ tiene asociada la matriz Hessiana

$$Hf_{(1,0)} = \begin{pmatrix} 0 & -1 \\ -1 & -2 \end{pmatrix}$$

y sus valores propios se calculan mediante la ecuación

$$0 = \begin{vmatrix} -\lambda & -1 \\ -1 & -2-\lambda \end{vmatrix} = \lambda^2 + 2\lambda - 1$$

que tiene soluciones

$$\lambda_1 = \frac{-2+\sqrt{8}}{2} > 0, \quad \lambda_2 = \frac{-2+\sqrt{8}}{2} < 0$$

lo qua hace también de $f(1,0) = 0$ un valor crítico silla.

d. Finalmente, el punto crítico $p_4 = (1/3, 1/3)$ tiene la matriz Hessiana

$$Hf_{(1/3,1/3)} = \begin{pmatrix} -2/3 & -1/3 \\ -1/3 & -2/3 \end{pmatrix}$$

y los valores propios se calculan al resolver la ecuación

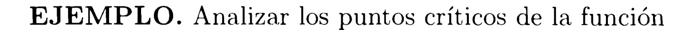
$$0 = \begin{vmatrix} -2/3-\lambda-1/3 & \\ -1/3 & \frac{-1}{3}-\lambda \end{vmatrix} = \lambda^2 + \frac{4}{3}\lambda + \frac{1}{3}$$

cuyas soluciones son

$$\lambda_1 = -1/3, \quad \lambda_2 = -1$$

que hace de $f\left(\frac{1}{3}, \frac{1}{3}\right) = 1/27$ un valor crítico mínimo.

Nuevamente, en este ejemplo nos hemos encontrado con puros puntos críticos no degenerados, pues todos los valores propios calculados son distintos de cero. ▷



EJEMPLO. Analizar los puntos críticos de la función

$$f(x,y) = y\sqrt{x} - y^2 - x + 6y$$

$\lhd$ Calculamos sus puntos críticos resolviendo el sistema,

$$\begin{cases} \frac{\partial f}{\partial x} = \frac{1}{2}yx^{-1/2} - 1 = 0 \\ \frac{\partial f}{\partial y} = x^{1/2} - 2y + 6 = 0 \end{cases}$$

que equivale a (si $x \neq 0$),

$$\begin{cases} \frac{y}{2\sqrt{x}} = 1 \\ \sqrt{x} = 2y - 6 \end{cases}$$

Al igualar las ecuaciones obtenemos $\frac{y}{2} = 2y - 6$, y resolviendo nos da $y = 4$.

Por lo tanto, obtenemos para x,

$$\frac{4}{2} = \sqrt{x}$$

lo que nos dice que $x = 4$.

De esta forma el único punto crítico de la función es $p = (4,4)$, y su valor crítico es $f(4,4) = 12$.

La matriz Hessiana viene dada por,

$$Hf_p = \begin{pmatrix} \frac{-y}{4\sqrt{x^3}} & \frac{1}{2\sqrt{x}} \\ \frac{1}{2\sqrt{x}} & -2 \end{pmatrix}_p$$

y de esta manera, la matriz Hessiana en el único punto crítico es

$$Hf_{(4,4)} = \begin{pmatrix} \frac{1}{64} & \frac{1}{4} \\ \frac{1}{4} & -2 \end{pmatrix}$$

cuyos valores propios se calculan mediante la solución de la ecuación

$$0 = \begin{vmatrix} \frac{1}{64} - \lambda & \frac{1}{4} \\ \frac{1}{4} & -2 - \lambda \end{vmatrix} = \lambda^2 + \frac{127}{64}\lambda - \frac{3}{32}$$

que al resolver nos da los valores de λ,

$$\lambda_1 = \frac{-127 + \sqrt{17665}}{128} > 0, \quad \lambda_2 = \frac{-127 - \sqrt{17665}}{128} < 0$$

lo que hace de $f(4,4) = 12$ un valor silla. Además, se tiene que el punto crítico es no degenerado. $\triangleright$

Hacemos hincapié de que el análisis de la gráfica de una función suave alrededor de un punto crítico no degenerado es simple y está totalmente descrito por la zoología de las formas cuadráticas básicas y el Lema de Sylvester. No obstante, cuando la matriz Hessiana en el punto crítico tiene un valor propio $\lambda = 0$, el análisis se complica y su estudio sale de nuestras posibilidades. La **Teoría de singularidades y Catástrofes** establecida por **R. Thom**, provee de la herramienta moderna necesaria para analizar el caso mencionado. A los puntos críticos tales que su matriz Hessiana tiene un valor propio nulo, se les llama **puntos críticos degenerados**.

Cabe mencionar que la Teoría expuesta aquí se extiende a varias variables como lo muestra el siguiente ejemplo.

EJEMPLO. Analizar los puntos críticos de la función

$$f(x,y,z) = x^2 + xy + y^2 + yz + z^2$$

$\triangleleft$ Claramente el dominio de la función f es toda la región espacial $\Omega = \mathbb{R}^3$.

Por otro lado, el gradiente de f está dado por

$$\nabla f_p = (2x + y, 2y + x + z, 2z + y)$$

de donde $\nabla f_p = (0,0,0)$, si y sólo sí, se cumple

$$\begin{cases} 2x + y = 0 \\ x + 2y + z = 0 \\ y + 2z = 0 \end{cases}$$

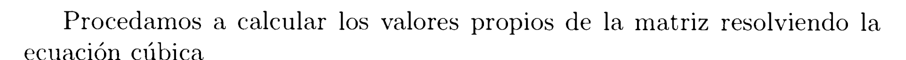
El determinante de tal sistema lineal es

$$\begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix} = 4 \neq 0$$

lo que nos dice que tal sistema sólo tiene la solución trivial $p = (0,0,0)$, que es entonces el único punto crítico.

La matriz Hessiana es, para este caso, la matriz simétrica de 3×3

$$B = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial z \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial z \partial y} \\ \frac{\partial^2 f}{\partial x \partial z} & \frac{\partial^2 f}{\partial y \partial z} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$



Procedamos a calcular los valores propios de la matriz resolviendo la ecuación cúbica

$$0 = det\left[\begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right] = \begin{vmatrix} 2-\lambda & 1 & 0 \\ 1 & 2-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{vmatrix}$$

$$= (2-\lambda)\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} - \begin{vmatrix} 1 & 1 \\ 0 & 2-\lambda \end{vmatrix} = (2-\lambda)[(2-\lambda)^2 - 1] - (2-\lambda)$$

$$= (2-\lambda)(\lambda^2 - 4\lambda + 2)$$

Las raíces de tal ecuación cúbica son los números positivos

$$\lambda_1 = 2, \quad \lambda_2 = 2 - \sqrt{2}, \quad \lambda_3 = 2 + \sqrt{2}$$

como se puede comprobar facilmente.

De esta manera, la matriz diagonalizada obtenida es

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2-\sqrt{2} & 0 \\ 0 & 0 & 2+\sqrt{2} \end{pmatrix}$$

con la diagonal en números positivos.

Esto haría que el valor crítico $f(0,0,0) = 0$ sea un valor mínimo local y que el punto crítico $p = (0,0,0)$ sea no degenerado. ▷

9.5 Multiplicadores de Lagrange

Damos inicio ahora al estudio de los valores extremos de una función que están sujetos a un dominio restringido dentro de su dominio natural. Para ello iniciaremos la discusión pensando que el dominio restringido mencionado tiene una característica adecuada en su forma: es un subconjunto del espacio, que es cerrado, y que está acotado (limitado) en el sentido que la norma de cada uno de sus puntos está superada por un número fijo apropiado.

DEFINICIÓN. Un conjunto $K \subset R^n$ que es cerrado y acotado se llamará **compacto**.

Hacemos una extensión del Teorema de Weierstrass (Reyes (1996)) para el caso de una función en varias variables, con un dominio **compacto**.

TEOREMA 9.1 (Weierstrass) *Toda función $f : K \to \mathbb{R}$ continua (con K compacto) alcanza sus valores máximo y su mínimo dentro de K.*



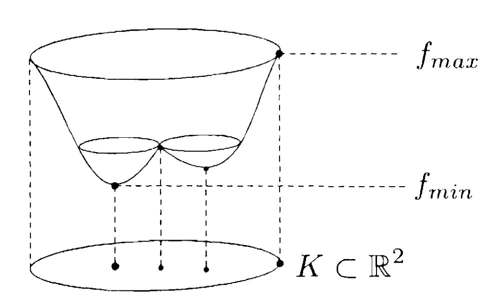


Figura 9.9: El Teorema de Weierstrass.

$\triangleleft$ Consideremos el caso en que la función f no sólo es continua, sino además es suave de clase C^r $(r \geq 1)$.

Si los puntos $p_1, p_2, \cdots, p_k$ son críticos para la función f en el interior de K, calculamos sus valores críticos

$$f(p_1), f(p_2), \cdots f(p_k)$$

Después, se buscan los extremos de f en los puntos $q_1, \cdots, q_l$ en la frontera, y en tales puntos obtenemos los valores

$$f(q_1),\ f(q_2), \cdots, f(q_l)$$

De entre el conjunto de todos los valores obtenidos

$$\{f(p_1), \cdots, f(p_k),\ f(q_1), \cdots, f(q_l)\}$$

se escogen el máximo y el mínimo (véase la figura 9.9). $\triangleright$

Por el Teorema de Weierstrass, para calcular los valores extremos de una función real de argumento vectorial definida en un conjunto compacto K, es necesario buscar en su interior los puntos críticos y evaluarlos, además de buscar en la frontera a aquellos puntos que pudieran dar valores extremos (véase la figura 9.9).

Procedamos a dar un método para localizar a tales puntos distinguidos en la frontera de K, suponiendo que tal frontera está definida como una superficie de nivel. A este método se le conoce como el de **multiplicadores de Lagrange**.

Sea S una superficie definida por una ecuación

$$g(x, y, z) = 0$$



donde g es una función suave en una región Ω que contiene a S y tal que

$$\nabla g_{(x,y,z)} \neq (0,0,0)$$

Dada $f : \Omega \to \mathbb{R}$ donde Ω es la región conteniendo a S, se dice que f tiene un **extremo restringido** a S en el punto $p \in S$, sí en una vecindad de $p \in S$ se cumple que en una vecindad de $p \in S$ se cumple que

$$f(p) \geq f(x,y,z)$$

en cuyo caso se dirá que $f(p)$ es un **máximo local** de f en S. Pero si en tal vecindad ocurre que

$$f(p) \leq f(x,y,z)$$

en este caso se dirá que $f(p)$ es un **mínimo local** de f en S.

El siguiente resultado debido a **Lagrange J.L.**, nos indica la forma de localizar a los puntos de S donde se encuentran los posibles extremos.

TEOREMA 9.2 (Lagrange) *Si f es un extremo en el punto $p \in S$ entonces existe un escalar λ (llamado el* **multiplicador de Lagrange***) tal que en el punto p se satisface la ecuación,*

$$\nabla f_p = \lambda \nabla g_p$$

$\lhd$ Sea $\gamma : J \to S$ una curva diferenciable pasando por $\gamma(0) = p$ y tomemos la composición

$$f \circ \gamma : J \to \mathbb{R}$$

dada por $(f \circ \gamma)(t) = f(\gamma(t))$.

Entonces, por la regla de la cadena, la composición es diferenciable y

$$\frac{d}{dt}(f \circ \gamma)(t) =< \nabla f_{\gamma(t)}, \dot{\gamma}(t) >$$

Al evaluar en el tiempo $t = 0$, se obtiene que

$$\frac{d}{dt}(f \circ \gamma)(0) =< \nabla f_p, \dot{\gamma}(0) >$$

Ya que $f(p)$ es extremo en S (es decir, un máximo o mínimo local sobre S) se tiene

$$0 = \frac{d}{dt}(f \circ \gamma)(0) =< \nabla f_p, \dot{\gamma}(0) >$$

debido a que para una función real de variable real, cuando alcanza un extremo (local o global) se tiene un punto crítico en $t = 0$.

Como $\dot{\gamma}(0)$ es un vector tangente en p y la curva γ es arbitraria, entonces esto nos dice que el vector ∇f_p es ortogonal al plano tangente T_pS.

Ya que también $\nabla g_p \perp T_pS$, tendremos entonces que el vector ∇f_p es paralelo a ∇g_p.

Por lo tanto , existe un número real λ diferente de cero tal que $\nabla f_p = \lambda \nabla g_p$ (véase la figura 9.10). $\triangleright$

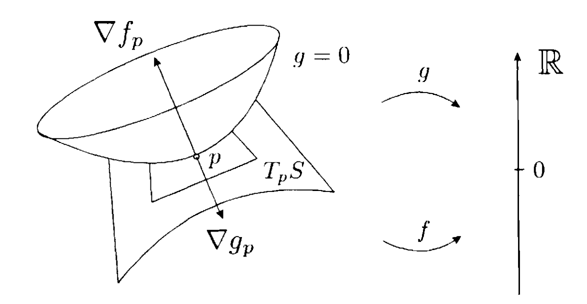


Figura 9.10: Multiplicador de Lagrange.

EJEMPLO. Hallar el máximo y el mínimo de la función

$$f(x, y) = x + y$$

sujetos a la restricción $x^2 + y^2 = 1$. En otras palabras, encontrar los valores máximo y mínimo de f sobre el círculo unitario S.

$\triangleleft$ Por el Teorema de Lagrange, es necesario resolver la ecuación

$$\nabla f_{(x,y)} = \lambda \nabla g_{(x,y)}$$

con $g(x, y) = x^2 + y^2 - 1$ definiendo a S como el conjunto de nivel

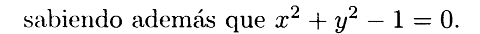
$$g^{-1}(0) = \{(x, y) |\, x^2 + y^2 - 1 = 0\}.$$

Al calcular los gradientes se tiene

$$\nabla f_{(x,y)} = (1, 1), \quad \text{y} \quad \nabla g_{(x,y)} = (2x, 2y)$$

y entonces, buscamos los valores de x, y y λ tal que satisfagan la ecuación

$$(1, 1) = \lambda(2x, 2y)$$



sabiendo además que $x^2 + y^2 - 1 = 0$.

Al resolver el sistema

$$\begin{cases} 1 = 2\lambda x \\ 1 = 2\lambda y \end{cases}$$

usando que $\lambda \neq 0$, se tiene que

$$\begin{cases} \frac{1}{2\lambda} = x \\ \frac{1}{2\lambda} = y \\ x^2 + y^2 = 1 \end{cases}$$

Por lo tanto,

$$\left(\frac{1}{2\lambda}\right)^2 + \left(\frac{1}{2\lambda}\right)^2 = 1$$

de donde

$$\frac{2}{4\lambda^2} = 1$$

es decir,

$$\lambda = \pm\sqrt{\frac{1}{2}} = \pm\frac{\sqrt{2}}{2}$$

De esta forma, obtenemos las soluciones de x, y en el círculo dadas por,

$$x = \frac{1}{2\left(\pm\frac{\sqrt{2}}{2}\right)} = \pm\frac{\sqrt{2}}{2}, \quad y = \pm\frac{\sqrt{2}}{2}$$

que implica que hay 4 puntos candidatos sobre el círculo $x^2 + y^2 = 1$, y que estos sean,

$$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \quad \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right), \quad \left(\frac{-\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \quad \left(\frac{-\sqrt{2}}{2}, \frac{-\sqrt{2}}{2}\right)$$

Si evaluamos cada uno de ellos se tiene que,

$$f\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} = 2\frac{\sqrt{2}}{2} = \sqrt{2}$$

$$f\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = 0, \quad f\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = 0, \quad f\left(\frac{-\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = -\sqrt{2}$$



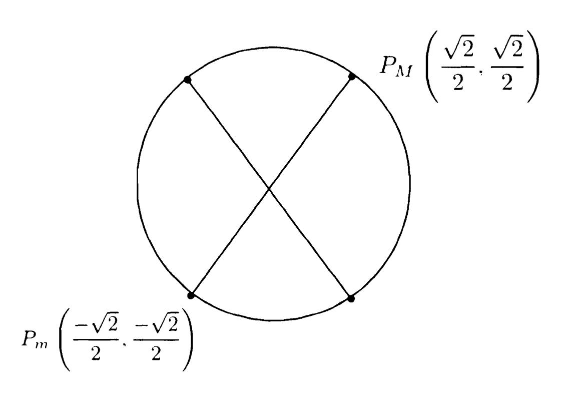


Figura 9.11: Extremos de $x + y$ en el círculo unitario.

Por lo tanto, los valores máximo y minimo de f sobre el círculo unitario son

$$f_{max} = \sqrt{2}, \quad \text{alcanzado con el punto} \quad p_M = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$$

$$f_{min} = -\sqrt{2}, \quad \text{alcanzado con el punto} \quad p_m = \left(\frac{-\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$$

La figura 9.11 ilustra este ejemplo. $\rhd$

EJEMPLO. Sea la función en dos variables,

$$f(x, y) = x + y^2$$

Calcular sus máximos y minimos sujetos a la restricción $2x^2 + y^2 = 1$.

$\lhd$ El gradiente de la función f se calcula,

$$\nabla f_p = (1, 2y)$$

y definimos la función auxiliar $g(x, y) = 2x^2 + y^2 - 1$ para la restricción.

Entonces, la curva plana $2x^2 + y^2 = 1$ que es una elipse, se puede escribir como la imagen inversa de 0 bajo g, es decir $g^{-1}(0)$ y además, sobre cada punto (x, y) de tal curva,

$$\nabla g(x, y) = (4x, 2y)$$

Ahora buscamos $\lambda \neq 0$ tal que $\nabla f_p = \lambda \nabla g_p$ para algunos puntos (x, y) sobre la curva, esto es, tal que

$$(1, 2y) = \lambda(4x, 2y)$$



o equivalentemente,

$$\begin{cases} 1 = 4\lambda x \\ 2y = 2\lambda y \end{cases}$$

sujetas a la condición de restricción $2x^2 + y^2 = 1$.

De la segunda ecuación, si $y \neq 0$ entonces $\lambda = 1$, y de la primer ecuación se tendría necesariamente que $1 = 4x$, lo que implicaría que $x = \frac{1}{4}$

Al usar la condición con $x = 1/4$, obtenemos,

$$2\left(\frac{1}{4}\right)^2 + y^2 = 1$$

que al resolver para y nos da,

$$y = \pm\sqrt{\frac{7}{8}}$$

Por esto, los puntos candidatos obtenidos mediante este análisis son,

$$\left(\frac{1}{4}, -\sqrt{\frac{7}{8}}\right), \quad \left(\frac{1}{4}, \sqrt{\frac{7}{8}}\right)$$

Consideremos ahora el caso cuando $y = 0$. Entonces, de la primer ecuación ($\lambda \neq 0$) se tiene que

$$\frac{1}{4\lambda} = x$$

Como que al sustituir en la ecuación de condición nos da (con $y = 0$),

$$2\left(\frac{1}{4\lambda}\right)^2 = 1$$

Tal ecuación nos provee de la solución para λ,

$$\lambda = \pm\sqrt{\frac{1}{8}}$$

y por lo tanto,

$$x = \frac{1}{4\left(\pm\sqrt{\frac{1}{8}}\right)} = \pm\frac{\sqrt{8}}{4} = \pm\frac{\sqrt{2}}{2}$$

De aquí, que los puntos candidatos sean en este caso,

$$\left(\frac{-\sqrt{2}}{2}, 0\right), \left(\frac{\sqrt{2}}{2}, 0\right)$$

Usando ahora el Teorema de Weierstrass, evaluamos todos los puntos obtenidos en ambos casos,

$$f\left(\frac{1}{4}, -\sqrt{\frac{7}{8}}\right) = \frac{1}{4} + \frac{7}{8} = \frac{9}{8}$$

$$f\left(-\frac{\sqrt{2}}{2}, 0\right) = -\frac{\sqrt{2}}{2} + 0^2 = \frac{-\sqrt{2}}{2}$$

$$f\left(\frac{1}{4}, \sqrt{\frac{7}{8}}\right) = \frac{1}{4} + \frac{7}{8} = \frac{9}{8}$$

$$f\left(\frac{\sqrt{2}}{2}, 0\right) = \frac{\sqrt{2}}{2} + 0^2 = \frac{\sqrt{2}}{2}$$

y concluimos que, sobre la elipse $2x^2 + y^2 = 1$, la función f tiene los extremos

$$f_{max} = \frac{9}{8}, \quad f_{min} = \frac{-\sqrt{2}}{2} \quad \triangleright$$

Damos ahora un ejemplo en tres variables donde se conjugan el Teorema de Weirestrass y el método de los multiplicadores de Lagrange.

EJEMPLO. Sea la función

$$f(x, y, z) = x^2 + xy + y^2 + yz + z^2$$

Hallar sus valores máximo y mínimo dentro de la bola de radio 1 con centro en el origen de coordenadas.

$\triangleleft$ Ya que la bola de radio 1 con centro en el origen es un cerrado y acotado, (**compacto**) utilizamos el Teorema de Weirestrass.

Este nos indica que hay que buscar los extremos en el interior $x^2 + y^2 + z^2 < 1$ a través de los puntos críticos , y en la frontera $x^2 + y^2 + z^2 = 1$ mediante el método de multiplicadores de Lagrange.

1. Los puntos críticos de la función

$$f(x, y, z) = x^2 + xy + y^2 + yz + z^2$$

se han localizado ya en un ejemplo previo (de hecho en todo $\mathbb{R}^3$), obteniendo apenas el punto crítico no degenerado $q = (0,0,0)$ con valor crítico $f(0,0,0) = 0$. Esto termina la búsqueda en el interior de la bola.

2. Procedamos entonces a buscar los extremos en la frontera $x^2 + y^2 + z^2 = 1$, es decir, los extremos de f sujetos a esta última restricción.

La frontera de la bola, que es una esfera de radio 1, está definida por la ecuación

$$x^2 + y^2 + z^2 = 1$$

Al calcular el vector gradiente de f se tiene entonces,

$$\nabla f_p = (2x + y,\, x + 2y + z, y + 2z)$$

Ahora sea la función auxiliar $g(x, y, z) = x^2 + y^2 + z^2 - 1$, entonces la esfera está definida como el conjunto $g^{-1}(0)$, y en cada punto,

$$\nabla g_p = (2x, 2y, 2z)$$

Buscamos entonces un escalar λ tal que $\nabla f_p = \lambda \nabla g_p$, es decir tal que

$$(2x + y, x + 2y + z, y + 2z) = \lambda(2x, 2y, 2z)$$

o equivalentemente que satisfaga las ecuaciones del sistema,

$$\begin{cases} 2x + y = 2\lambda x \\ \lambda + 2y + z = 2\lambda y \\ y + 2z = 2\lambda z \end{cases}$$

En otras palabras, buscamos una condición para λ tal que el sistema

$$\begin{cases} 2(1-\lambda)x + y = 0 \\ x + 2(1-\lambda)y + z = 0 \\ y + 2(1-\lambda)z = 0 \end{cases}$$

tenga soluciones no triviales. Pero, el determinante de este sistema homogéneo se anula cuando

$$D = \begin{vmatrix} 2(1-\lambda) & 1 & 0 \\ 1 & 2(1-\lambda) & 1 \\ 0 & 1 & 2(1-\lambda) \end{vmatrix} = 0 \Longleftrightarrow (1-\lambda)[4(1-\lambda)^2 - 2] = 0$$

lo que nos dice que hay soluciones no triviales para el sistema sólo para valores del multiplicador λ dados por

$$\lambda = 1, \quad \lambda_{2,3} = 1 \pm \frac{\sqrt{2}}{2}$$



Buscamos, consecuentemente, las soluciones no triviales considerando el sistema lineal dado por la pareja,

$$\begin{cases} 2(1-\lambda)x + y = 0 \\ x + 2(1-\lambda)y + z = 0 \end{cases}$$

y para valores que este sistema tenga soluciones no triviales se necesita que

$$0 \neq \begin{vmatrix} 2(1-\lambda) & 1 \\ 1 & 2(1-\lambda) \end{vmatrix}$$

Pero tal determinante se anula, sí y sólo sí,

$$4(1-\lambda)^2 - 1 = 0$$

es decir, $\lambda = \frac{1}{2}, 2/3$.

Por lo tanto si $\lambda \neq \frac{1}{2}, \frac{3}{2}$, entonces

$$D = \begin{vmatrix} 2(1-\lambda) & 1 \\ 1 & 2(1-\lambda) \end{vmatrix} \neq 0$$

y el sistema lineal inicial tiene soluciones no triviales, que a continución conseguimos.

Si hacemos $z = t$, entonces para el sistema con parámetro t,

$$\begin{cases} 2(1-\lambda)x + y = 0 \\ x + 2(1-\lambda)y = -t \end{cases}$$

se tienen los determinantes

$$D_z = \begin{vmatrix} 0 & 1 \\ -t & 2(1-\lambda) \end{vmatrix} = t, \quad D_y = \begin{vmatrix} 2(1-\lambda) & 0 \\ 1 & -t \end{vmatrix} = -2t(1-\lambda)$$

y por lo tanto la solución en general es la recta,

$$\begin{cases} x = \frac{t}{4(1-\lambda)^2-1} \\ y = \frac{-2t(1-\lambda)}{4(1-\lambda)^2-1} \\ z = t \end{cases}$$

que depende de los valores encontrados de λ.

Ya que tenemos las soluciones del sistema, utilizamos las condiciones de los multiplicadores obtenidos $\lambda = 1, \lambda_2, \lambda_3$.

a. Multiplicador $\lambda = 1$.



Se obtiene en este caso la ecuación de la recta

$$\begin{cases} x = -t \\ y = 0 \\ z = t \end{cases}$$

y además $x^2 + y^2 + z^2 = 1$, de donde $t^2 + t^2 = 1$, lo que implica que.

$$t = \pm\sqrt{\frac{1}{2}} = \pm\frac{\sqrt{2}}{2}$$

Por lo tanto sustituyendo el valor de t en la ecuación de la recta se tienen los puntos,

$$p_1 = \left(-\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right), \; p_2 = \left(\frac{\sqrt{2}}{2}, 0, \frac{-\sqrt{2}}{2}\right)$$

b. Multiplicador $\lambda = 1 - \frac{\sqrt{2}}{2}$.

Se obtiene en este caso la ecuación de la recta dada por,

$$\begin{cases} x = t \\ y = -\sqrt{2}t \\ z = t \end{cases}$$

condicionada a $x^2 + y^2 + z^2 = 1$, de donde,

$$t^2 + 2t^2 + t^2 = 1 \iff t^2 = \frac{1}{1}$$

lo que implica que $t = \pm\frac{1}{2}$.

De esta manera, se tienen entonces los puntos

$$p_3 = \left(\frac{1}{2}, \frac{-\sqrt{2}}{2}, \frac{1}{2}\right), \quad p_4 = \left(-\frac{1}{2}, \frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$$

c. Multiplicador $\lambda = 1 + \frac{\sqrt{2}}{2}$

Se tiene, en este caso, el sistema solución

$$\begin{cases} x = t \\ y = \sqrt{2}t \\ z = t \end{cases}$$

y además $x^2 + y^2 + z^2 = 1$, de donde $t^2 + 2t^2 + t = 1$, lo que implica que $t = \pm\frac{1}{2}$.

Los puntos obtenidos son, en este caso,

$$p_5 = \left(\frac{1}{2}, \frac{\sqrt{2}}{2}, \frac{1}{2}\right), \quad p_6 = \left(-\frac{1}{2}, \frac{-\sqrt{2}}{2}, -\frac{1}{2}\right)$$

Se han obtenido mediante todo el proceso los puntos candidatos del Teorema de Weierstrass,

$$\left(-\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right), \left(\frac{\sqrt{2}}{2}, 0, -\frac{\sqrt{2}}{2}\right) \left(\frac{1}{2,} - \frac{\sqrt{2}}{2}, \frac{1}{2}\right),$$

$$\left(-\frac{1}{2}, \frac{\sqrt{2}}{2}, -\frac{1}{2}\right), \left(\frac{1}{2}, \frac{\sqrt{2}}{2}, \frac{1}{2}\right), \left(-\frac{1}{2}, -\frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$$

Al evaluar tales puntos obtenemos sus valores correspondientes

$$f\left(-\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right) = 1$$

$$f\left(-\frac{\sqrt{2}}{2}, 0, -\frac{\sqrt{2}}{2}\right) = 1$$

$$f\left(\frac{1}{2}, -\frac{\sqrt{2}}{2}, \frac{1}{2}\right) = 1 - \frac{\sqrt{2}}{4} - \frac{\sqrt{2}}{4} = 1 - \frac{\sqrt{2}}{2}$$

$$f\left(-\frac{1}{2}, \frac{\sqrt{2}}{2}, -\frac{1}{2}\right) = 1 - \frac{\sqrt{2}}{4} - \frac{\sqrt{2}}{4} = 1 - \frac{\sqrt{2}}{2}$$

$$f\left(\frac{1}{2}, \frac{\sqrt{2}}{2}, \frac{1}{2}\right) = 1 + \frac{\sqrt{2}}{4} + \frac{\sqrt{2}}{4} = 1 + \frac{\sqrt{2}}{2}$$

$$f\left(-\frac{1}{2}, -\frac{\sqrt{2}}{2}, -\frac{1}{2}\right) = 1 + \frac{\sqrt{2}}{4} + \frac{\sqrt{2}}{4} = 1 + \frac{\sqrt{2}}{2}$$

De esta manera, sobre la frontera (esfera) se tienen los valores máximo y mínimo dados por

$$f_{min} = 1 - \frac{\sqrt{2}}{2}, \quad f_{max} = 1 + \frac{\sqrt{2}}{2}$$

Consecuentemente, ya que el interior se tenía que $f(0,0,0) = 0$, entonces los valores máximo y mínimo de f en toda la bola cerrada son,

$$f_{min} = 0 \quad \text{y} \quad f_{max} = 1 + \frac{\sqrt{2}}{2}. \quad \triangleright$$



Ejercicios

1. Calcular los valores extremos locales de cada función dada, utilizando para ello el Lema de Sylvester. Clasificarlos como máximos, mínimos o sillas.

a. $f(x,y) = xy^2(1-x-y)$

b. $f(x,y) = x^2 + 4y^2 - 4x$

c. $f(x,y) = (x^2+y^2)(1-e^{(-x^2+y^2)})$

d. $f(x,y) = e^x \sin y$

e. $f(x,y) = x^3 + y^3 - 15xy$

2. a. Dar los valores extremos de la función de dos variables $f(x,y) = x^2 + y^2 - xy - 4y$ en la región limitada por las rectas $x = 0, y = 0, \quad x + y - 1 = 0$

b. Dar los valores máximo y mínimo de la función $f(x,y) = x^2 + 3y^2 + x - y$ en el triángulo limitado por las rectas $x = 1$, $y = 1$, $x + y = 1$

3. Dar los valores extremos de la función dada, con la restricción que se indica.

a. $f(x,y,z) = 2x^2 + 2y^2 - z^2, \quad x - y + z = 0$

b. $f(x,y,z) = ze^{xy}, \quad x^2 + y^2 = z$

4. a. Calcular los valores máximo y mínimo de la función $f(x,y) = x + y^2$ en el círculo unitario.

b. Calcular el valor máximo y el valor mínimo de la función $f(x,y) = x + y + z$ sobre la esfera $x^2 + y^2 + z^2 = 1$.

5. a. Calcular los valores extremos de la función $f(x,y,z) = x + 2y - 2z$ en la bola unitaria $x^2 + y^2 + z^2 \leq 1$.

b. Calcular el valor máximo y el valor mínimo de la función $f(x,z) = x - y - z$ dentro de la bola $x^2 + y^2 + z^2 \leq 1$.

6. Sea un círculo de radio R. Dar las dimensiones del triángulo que se puede inscribir dentro del círculo, y que tenga área máxima.

7. Hallar las dimensiones de un paralelepípedo de volumen máximo que tenga fija una área dada A.

8. De entre los rectángulos con área dada A, hallar las dimensiones de aquél que tiene perimetro con valor mínimo.

9. Dar la distancia más corta de la elipse $\frac{x^2}{4} + y^2 = 1$ hasta la recta $x + y - 4 = 0$.


Casa abierta al tiempo

Capítulo 10

Funciones implícitas

Sea $\Omega \subset \mathbb{R}^n$ una región con frontera y sea $f : \Omega \to f(\Omega)$ un **difeomorfismo global** , esto es, una función diferenciable, invertible con una inversa diferenciable $g : f(\Omega) \to \Omega$. El siguiente resultado cuya prueba se omite (pues requiere de las herramientas de la teoría de la **Topología Algebraica**) nos indica que los puntos interiores de Ω bajo la aplicación de f son llevados al interior de $f(\Omega)$, y los puntos de la frontera de Ω son llevados en la frontera de $f(\Omega)$. (Véase la figura 10.1).

TEOREMA 10.1 (de la invarianza del dominio) *Sean $\Omega \subset \mathbb{R}^n$ una región con frontera $\partial\Omega$ y sea $f : \Omega \to f(\Omega) \subset \mathbb{R}^n$ un difeomorfismo global. Entonces $f(\Omega)$ es una región con frontera en $\mathbb{R}^n$. Los puntos interiores de Ω son llevados bajo f en puntos interiores de $f(\Omega)$ y los puntos frontera en $\partial\Omega$ son llevados bajo f en la frontera $\partial f(\Omega)$.*

EJEMPLO. Sea la función $\mathbb{R}^2 \to \mathbb{R}^2$ dada por

$$f(x, y) = (ax, by) = (u, v), \quad \text{con} \quad a, b > 0.$$

$\lhd$ Si se considera la región circular con frontera conformada por el disco unitario de radio 1,

$$\Omega = \left\{(x, y) |\, x^2 + y^2 \leq 1\right\}$$

se tienen que bajo la función f la región Ω es mapeado en la región elipsoidal con frontera

$$f(\Omega) = \left\{(u, v) |\, \frac{u^2}{a^2} + \frac{v^2}{b^2} \leq 1\right\}$$

De las ecuaciones,

$$\begin{cases} u = ax \\ v = by \end{cases}$$

se obtienen las coordenadas inversas,

$$\begin{cases} x = \frac{u}{a} \\ y = \frac{v}{b} \end{cases}$$

que dan una función inversa $\mathbb{R}^2 \to \mathbb{R}^2$ dada por,

$$g(u, v) = \left(\frac{u}{a}, \frac{v}{b}\right) = (x, y)$$

Claramente la función $f : \Omega \to f(\Omega)$ es **difeomorfismo** que lleva el interior del disco Ω en el interior del elipsoidal $f(\Omega)$, llevando la frontera del disco $x^2 + y^2 = 1$ en la elipse $\frac{u^2}{a^2} + \frac{v^2}{b^2} = 1$, cumpliéndose el Teorema de la invarianza del dominio, (véase la figura 10.1).

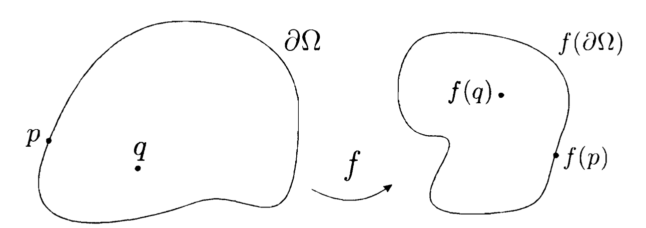


Figura 10.1: Teorema de invarianza del dominio.

EJEMPLO. Consideremos la transformación $\mathbb{R}^2 \to \mathbb{R}^2$ dada por

$$f(r, \varphi) = (r \cos \varphi, r \operatorname{sen} \varphi)$$

donde un punto $p = (r, \varphi)$ toma unidades en números positivos para la primer coordenada y en radianes para la segunda.

$\lhd$ Sea Ω la región con frontera en el plano $\mathbb{R}^2$ dada por el cuadrado

$$\Omega = \{(r, \varphi) | \quad 0 \leq r \leq 1, \quad 0 \leq \varphi \leq \pi/4\}$$

entonces, de la definición d de la función f es claro que en coordenadas x, y del codominio, las relaciones se escriben

$$x = r \cos \varphi, \quad y = r \operatorname{sen} \varphi$$

que es el cambio de coordenadas cartesiano por **coordenadas polares**.

Una simple inspección nos muestra que la región $f(\Omega)$ es la cuarta parte del círculo unitario en el primer cuadrante, con la frontera incluida. Aquí nuevamente los puntos interiores del cuadrado Ω son llevados en los interiores de $f(\Omega)$ bajo la función f, y la frontera de Ω en la frontera de



$f(\Omega)$, cumpliéndose totalmente el **Teorema de invarianza de dominio**. $\triangleright$

Ahora damos paso a la demostración de un resultado que es equivalente al Teorema de la función inversa, y que en virtud de que ya se conoce éste, se presenta como una de sus consecuencias: El **Teorema de la función implícita**. Iniciamos la discusión con un ejemplo.

EJEMPLO. Sea la función en dos variables

$$f(x,y) = x^2 + y^2$$

$\triangleleft$ Si $c \in \mathbb{R}^+ \cup \{0\}$ es un valor arbitrario, se tiene que la curva de nivel $f^{-1}(c)$ está dada por la ecuación $x^2+y^2 = c$. Sea $(x,y) \in f^{-1}(c)$, entonces al despejar la variable y se tiene que $y = \pm\sqrt{c - x^2}$.

Si y es una coordenada positiva, entonces necesariamente $y = \sqrt{c - x^2}$.

De esta manera la variable y se puede escribir localmente como una función $y = h(x)$ en los puntos $(x,y) \in f^{-1}(c)$, es decir para tales puntos,

$$c = f(x,y) = f(x,h(x))$$

la función $y = h(x)$ se llama una **función implícita** para la función f.

Observamos que la derivada implícita y' se calcula mediante la derivación de la igualdad $x^2 + y^2 = c$, es decir,

$$2x + 2yy' = 0$$

o equivalentemente,

$$y' = \frac{-2x}{2y} = \frac{x}{y} = \frac{-x}{h(x)} = \frac{-x}{\sqrt{c - x^2}}$$

la cual tiene sentido para $-\sqrt{c} < x < \sqrt{c}$.

Por ejemplo si $x = 0$, entonces $y = \sqrt{c}$ de donde en el punto $(0, \sqrt{c})$ se tendría una pendiente $y'(0) = 0$. $\triangleright$

Enunciamos y demostramos el Teorema de la función implicita para dos variables. La generalización se muestra apenas en los ejemplos subsecuentes.

TEOREMA 10.2 (de la Función Implícita). *Sea $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ una función de clase C^r y sea $p = (a,b) \in \Omega$ tal que $f(a,b) = c$ y tal que $\frac{\partial f}{\partial y}(a,b) \neq 0$. Entonces existe una función $y = h(x)$ de clase C^r definida en una vecindad del punto $x = a$ en $\mathbb{R}$, tal que $h(a) = b$ y todos los puntos*



de la gráfica de h están contenidos en el conjunto de nivel $f^{-1}(c)$, es decir, sí el punto $(x, y) = (x, h(x))$ está en la gráfica de h en el plano, entonces,

$$c = f(x, y) = f(x, h(x))$$

◁ Sea la función definida por $\phi(x, y) = (x, f(x, y))$. Se tiene

$$D\phi_p = \begin{pmatrix} 1 & 0 \\ f_x & f_y \end{pmatrix}_p$$

donde f_x es la derivada parcial de f en la variable x, y f_y la derivada parcial de f respecto a la variable y. De esto, el Jacobiano es

$$J\phi_p = f_y = \frac{\partial f}{\partial y}(p) \neq 0$$

y por el Teorema de la función inversa, existen, una vecindad de p en Ω y una vecindad de $q = f(p)$ donde está definida para la función dada ϕ, una función inversa local ψ de clase C^r.

Sean las coordenadas (x, z) en el codominio de ϕ y póngase

$$\psi(x, z) = (x, g(x, z))$$

Si además se denotan $z = f(x, y)$ y $y = g(x, z)$, entonces

$$(x, y) = \psi \circ \phi(x, y) = \psi(\phi(x, y)) = \psi(x, f(x, y))$$

$$(x, z) = \phi \circ \psi(x, z) = \phi(\psi(x, z)) = \phi(x, g(x, z))$$

Definamos entonces para $c = f(a, b)$, la función suave de clase C^r, en la variable x en una vecindad de $x = a$, dada por

$$h(x) = g(x, c)$$

se tiene entonces

$$\phi(x, h(x)) = \phi(x, g(x, c)) = \phi(\psi(x, c)) = (x, c)$$

y por otro lado, para x en la vecindad de $x = a$,

$$\phi(x, h(x)) = (x, f(x, h(x)))$$

de donde necesariamente $f(x, h(x)) = c$.

Ya que $c = f(a, b)$ y $f(a, h(a)) = c$, se tiene que $h(a) = b$ lo que termina la prueba del Teorema (véase el la figura 10.2). ▷



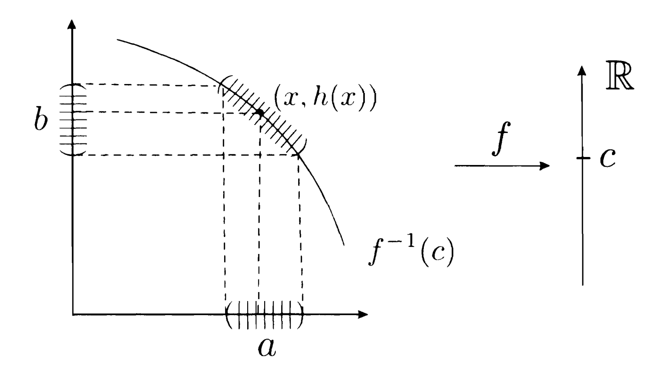


Figura 10.2: Teorema de la función implícita.

Observemos que para este caso, sí sobre la curva de nivel $f(x,y) = c$ la variable y se puede escribir implícitamente por la función $y = h(x)$ en la vecindad del punto $(x = a, y = b)$, entonces al derivar parcialmente la ecuación de la curva de nivel $f(x,y) = c$ respecto a la variable x se obtiene la derivada implícita ordinaria $y' = h'(x)$ respecto a x.

◁ Si $f(x,y) = c$, entonces, por la regla de la cadena,

$$\frac{\partial f}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} = 0$$

o equivalentemente, sabiendo que $y = h(x)$

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}y' = 0$$

que se puede escribir $f_x + f_y y' = 0$.

De esta manera,

$$y' = \frac{-f_x}{f_y}$$

sabiendo que $f_y = \frac{\partial f}{\partial y} \neq 0$. ▷

EJEMPLO. Sea la función en dos variables dada por,

$$f(x,y) = x\cos xy$$

◁ Al calcular la derivada parcial respecto a la variable y se tiene

$$f_y(x,y) = -x^2 \operatorname{sen} xy$$

y particularmente en el punto $p = (1, \pi/2)$, obtenemos

$$f_y(1, \pi/2) = 1 \neq 0$$

De aquí, por el Teorema de la función implícita, existe una función $y = h(x)$ en la cercanía de $x = 1$, tal que $h(1) = \pi/2$ y tal que la derivada de tal función en $x = 1$ se calcula

$$y'(1) = \frac{-f_x}{f_y}|_{(1,\pi/2)} = \frac{\cos xy - xy \operatorname{sen} xy}{-x^2 \operatorname{sen} xy}|_{(1,\pi/2)} = -\frac{\pi}{2}$$

Procedamos a calcular en este caso la función implícita resolviendo el caso general para las curvas de nivel. Si $c \in \mathbb{R}$, entonces,

$$f(x, y) = x \cos xy = c$$

implica que, para $x \neq 0$, se tiene

$$\cos xy = \frac{c}{x}$$

o equivalentemente, al despejar la variable y obtenemos

$$y = \frac{1}{x} \arccos\left(\frac{c}{x}\right)$$

Así, en $x = 1$, $y = \pi/2$ se cumple $c = f(1, \pi/2) = 0$, de donde, la constante $c = 0$ y la función particular buscada es

$$y = h(x) = \frac{1}{x} \arccos(0) = \frac{\pi/2}{x}$$

cuya derivada es,

$$y' = h'(x) = \frac{-\pi/2}{x^2}$$

De esta manera, $y = h(1) = \pi/2$ y $y' = h'(1) = -\pi/2$ lo que comprueba nuestros cálculos obtenidos. ▷

El **Teorema de la función implícita** para el caso de funciones de tres variables se establece de la siguiente manera, y omitimos su prueba que procede de igual forma que para el caso de dos variables.

TEOREMA 10.3 (Función implícita). *Sea $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ una función de clase C^r y sea $p = (a, b, c)$ un punto arbitrario de Ω tal que $f(a, b, c) = d$ y $\frac{\partial f}{\partial z}(a, b, c) \neq 0$. Entonces en una vecindad del punto (a, b) en el plano, existe una función $z = h(x, y)$ de clase C^r tal que la gráfica*



está contenida en la superficie de nivel $f(x, y, z) = d$, esto es, para puntos (x, y) cercanos del punto (a, b) se tiene que

$$f((x, y), h(x, y)) = d \quad y \quad h(a, b) = c.$$

Más aún se cumplen las igualdades para las parciales,

$$\frac{\partial h}{\partial x}(a, b) = \frac{-f_x}{f_z}(a, b, c)$$

$$\frac{\partial h}{\partial y}(a, b) = \frac{-f_y}{f_z}(a, b, c)$$

El siguiente ejemplo muestra que en general no se puede calcular explícitamente la función h, sino que apenas se demuestra su existencia.

EJEMPLO. Sea la función

$$f(x, y, z) = z^3 - z - xy \operatorname{sen} z$$

Mostrar que en una vecindad del punto $(0, 0, 0)$ la variable z se puede poner en función de las otras dos variables sobre la superficie de nivel

$$f(x, y, z) = f(0, 0, 0) = 0$$

◁ Calculamos la derivada parcial de f respecto a z en el punto $(0, 0, 0)$, obteniendo

$$\frac{\partial f}{\partial z}(0, 0, 0) = 3z^2 - 1 - xy \cos z|_{(0,0,0)} = -1 \neq 0$$

y por el Teorema de la función implicita 10.3, existe en una vecindad del punto $(0, 0)$ en el plano, y una función $z = h(x, y)$ de clase C^r tal que,

$$h(0, 0) = 0 \quad \text{y} \quad f(x, y, h(x, y)) = f(0, 0, 0) = 0$$

Por otro lado, las derivadas parciales de h en el punto $(0, 0)$ se calculan

$$\frac{\partial h}{\partial x}(0, 0) = \frac{-f_x}{f_z}(0, 0, 0) = \frac{-y \operatorname{sen} z}{3z^2 - 1 - xy \cos z}(0, 0, 0) = 0$$

$$\frac{\partial h}{\partial y}(0, 0) = \frac{-f_y}{f_z}(0, 0, 0) = \frac{-x \operatorname{sen} z}{3z^2 - 1 - xy \cos z}(0, 0, 0) = 0$$

Al tratar de proceder a buscar la función $z = h(x, y)$ a partir de la superficie de nivel $z^3 - z - xy \operatorname{sen} z = 0$, podemos observar que no es simple despejar a z en términos de las otras variables (x, y), lo que nos dice que, en general, la función h mencionada en el Teorema 10.3 no es fácil de encontrar. ▷



Ejercicios.

1. Mostrar que $f(x,y)=c$ define una función implícita $y=h(x)$ alrededor del punto dado $p=(a,b)$. Calcular $y'=h'(a)$.

a. $f(x,y)=y\cos xy$ en $p=(\pi/2,1)$
b. $f(x,y)=xe^y-y+4$ en $p=(-2,1)$
c. $f(x,y)=xy+y^4+x^4$ en $p=(1,-1)$
d. $f(x,y)=x^2+2xy+y^2$ en $p=(1,1)$

2. Mostrar que la función $f(x,y,z)=c$ define una función implícita $z=h(x,y)$ en una vecindad del punto dado $p=(a,b,c)$. Calcular $\frac{\partial h}{\partial x}(a,b)$ y $\frac{\partial h}{\partial y}(a,b)$.

a. $f(x,y,z)=z-z^3+xy\operatorname{sen} z$ en $p=(0,0,\pi/2)$
b. $f(x,y,z)=x^2+y^2+z^2+2xyz$ en $p=(1,1,1)$
c. $f(x,y,z)=x-y+2z-e^{xyz}$ en $p=(0,1,1)$

Índice





Bibliografía

Anton, H., Introducción al Álgebra Lineal, ed. Limusa, México, 1981.

Lima, E.L., Curso de Análise, Projeto Euclídes, IMPA, Rio de Janeiro, Brasil, 1992.

Reyes, J. G., Cálculo Diferencial para las Ciencias Naturales, Trillas, 1996.


Casa abierta al tiempo



CÁLCULO DIFERENCIAL EN VARIAS VARIABLES, se terminó de imprimir en el mes de diciembre de 2002 en la Sección de Talleres Gráficos del Departamento de Publicaciones de Rectoría General de la Universidad Autónoma Metropolitana. Se tiraron 1000 ejemplares sobre papel gráfico bond de 44.5 kgs. en tipos Times de 10 y 11 pts. y Times bold de 14 pts. Diseño de portada a cargo de D.C.G. Ma. de Lourdes Pérez Granados de la Sección de Textos y Formación del mismo Departamento..


Casa abierta al tiempo


ISBN: 970-31-0096-1
9 789703 100965